DOCUMENT RESUME

ED 068 498 SP 005 919

AUTHOR Baker, Eva L.
TITLE Preparing Instructional Materials for Educational Developers. Final Report.
SPONS AGENCY National Center for Educational Research and Development (DHEW/OE), Washington, D.C.
BUREAU NO BR-1-0327
PUB DATE Mar 72
GRANT OEG-0-71-0645
NOTE 97p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Educational Development; *Instructional Innovation; *Instructional Materials; *Instructional Programs; *Research Design; Teaching Techniques

ABSTRACT

The report describes the development of a package of training materials to enable instructional product developers to prepare first-draft materials. Five instructional techniques, derived from laboratory research and selected through a survey of eminent instructional psychologists, comprised the substance of the materials. The techniques were direct practice, knowledge of results, prompting, task description, and control or inspection behavior. Eight developmental and field tests were conducted on the materials using subjects drawn from universities, schools and industry. Results of the field test indicated that the materials were replicable in their instructional process and effects. Controlled variation studies are also reported. (Related document is SP 005 918.) (Author/MJM)

Final Report

Project No. 1-0027
Grant No. OEG-0-71-0645

Eva L. Baker
Department of Education
University of California
405 Hilgard Avenue
Los Angeles, California 90024

PREPARING INSTRUCTIONAL MATERIALS FOR EDUCATIONAL DEVELOPERS

March 1972

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Office of Education

National Center for Educational Research and Development

Final Report

Project No. 1-0027
Grant No. OEG-0-71-0645

PREPARING INSTRUCTIONAL MATERIALS FOR EDUCATIONAL DEVELOPERS

Eva L. Baker

Department of Education
University of California
405 Hilgard Avenue
Los Angeles, California 90024

March 1972

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education

National Center for Educational Research and Development

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

AUTHOR'S ABSTRACT

The report describes the development of a replicable package of training materials to enable instructional product developers to prepare first-draft materials. The objective of the materials was as follows:

> Given a set of specifications for short-term instructional sequences, the trainee will be able to produce a first-draft of instruction that includes the techniques described in the materials.

Five instructional techniques, derived from laboratory research and selected through a survey of eminent instructional psychologists, comprised the substance of the materials. The techniques were direct practice, knowledge of results, prompting, task description and control of inspection behavior. The final package required from 12-15 hours to complete and consisted of texts, practice discrimination sequences, models of instruction, revision and writing simulations, an annotated bibliography and criterion measures. The report describes the procedures employed to develop each of these components.

Eight developmental and field tests were conducted on the materials, using subjects drawn from universities, schools and industry. Results of the field tests indicated that the materials were replicable in their instructional process and effects. In addition, a series of controlled variation studies were conducted during the development process for the purpose of providing input to project decision-making. These studies are reported in detail.

INTRODUCTION

Two visible and conflicting alternatives exist for solving the major problems of instruction in the schools. The first set of solutions depends upon analysis and application of scientific findings and procedures. The translation of these findings and procedures results in a form of technology, and many prominent figures in the field of instructional psychology advocate such an approach (Glaser, 1966, Gagne, 1969; Schutz, 1970; Lumsdaine, 1964). The thrust of a second alternative solution to educational problems centers not upon the application of systematic procedures but rather on the redefinition of the proper scope of instructional endeavor from something measurable and, hence, verifiable to a more subjective, personal activity. Many of the proponents of less structured schooling come to their position because of frustration and disillusionment with their instructional experiences, and out of such distress springs their pattern for educational improvement. Instructional technologists, on the other hand, operate from a base of optimism. They assume that most problems can be solved; most phenomena can be explained. The onus is upon the technologist to select or to discover the solution. It was out of such optimism that this project emerged.

One of the few common agreements in educational thought is that change is a complex process. A dictum to employ a particular innovation or procedure inevitably is ignored. The history of instructional innovation is a dark and checkered one, and rarely have ideas disseminated through teacher education institutions, for instance, demonstrated positive impact on the achievement of the students.

In the last six years, however, the reliance for the dissemination of "better" approaches has shifted from teacher training to those agencies which develop instructional materials and procedures, sometimes described as systems. Such systems are designed to produce achievement of particular competencies in their users. These materials and procedures are planned to be reproducible and to follow the basic tenets of empirical development. The growing success and acceptance of early materials developed according to empirical procedures have encouraged educators to devote more attention to additional instructional development endeavors. Unfortunately, this interest has not been matched by an expanding force of personnel with training and expertise in the process of instructional design. The recognition of the need for suitably trained research and development individuals has contributed to the funding of the graduate research training program. However, the unquestionably localized, if not insular, consequences of university-based training programs precludes the widespread distribution of training. At the same time, the doctoral degree, heretofore the passport to the research realm, is being questioned as a prerequisite for effective work in instructional development. One solution to the troublesome problem of training development personnel involves the application of product development techniques to the training mission. Replicable materials which could be successfully used in a variety of sites present likely means for ameliorating the training need. This project was developed, in part, to test the notion that training in skills relevant to instructional design could be mediated effectively, and primarily, by instructional materials.

PROCEDURES

Project Selection

A survey of materials available to train instructional developers revealed certain areas of deficiency. For instance, attention has been given to the development of specifications (Mager, 1962, Popham, 1967) to the conduct of task analysis (Gagne, 1965; Briggs, 1970) and developmental testing (Markle, 1967) and field testing (Caliperg, 1969). However, materials relevant to the production of instructionally sound first drafts were less visible. Briggs (1970) devotes only a small portion of his book specifically to instruction and suggests that readers refer to books on "programming". The best such book, in the writer's opinion, has been written by Susan Markle (1970), however, the book is frankly a book on programming. While much of the content may be extrapolated to other instructional formats, the learner must do such translation for himself. A substantial portion of the book is directed to frame editing. Therefore, the perceived deficiency in replicable training materials devoted specifically to the development of skill in production of first draft instructional materials, motivated the decision to concentrate the project development effort on this specific area.

Overview of Work Plan

In a nonobvious display of naivete, the author was determined to conduct this project the "right" way, that is, in accordance with prescriptions for instructional development ventures. Thus, the project began with a delightful mixture of self-consciousness and smugness, both of which were soon to be dissipated. The project was planned to involve four major phases: (1) formulation, specification, and initial prototyping; (2) draft writing and developmental testing; (3) revision; and (4) demonstration testing. The first stage, where the substance of the project was designed to extend from December 1 to March 1. Phase two was scheduled for April through June; phase three in July; phase four was planned for August and September with the final report to be prepared in October. The planned schedule should be kept in mind as a description of the actual operation unfolds.

The staff members were also directed to keep track of the kinds and numbers of decisions made so that an adequate record of the development effort would be available.

Resume of Activities

Phase One: Formulation, Specification and Prototyping -- Dec. 1970-March 1971

Because the proposal committed the project to a set of gross specifications, the formulation phase of the project was relatively brief. Although the inclusion of five instructional techniques had been stipulated, the project staff needed first to determine which techniques would be those selected from the content of the materials.

During this period the following tasks were accomplished:

1. We conducted a survey of instructional psychologists to assess their preference for various instructional techniques.

2. Survey results were tabulated and presented in a paper at the Annual Meeting of the American Educational Research Association in New York.

3. Tentative selection of six instructional techniques was made: practice, prompting, knowledge of results, control of inspection behavior, task description and advance organizer. Selection was guided by both the results of the survey and the staff's attempt to render the techniques in instructional sequences.

4. Preliminary contacts were made for developmental and field test sites for spring and summer.

5. Examples of short-term instructional sequences in a variety of subject matters were developed. These sequences were to be suitable for different aged learners and to provide a sense of variety with which instruction might be produced. The sequences were planned to be used in discrimination and revision sections of the project.

6. Final specification of the objectives were generated and are presented below:

> Given a set of specifications from which to choose, the trainee would be able to develop an instructional sequence which exhibited the use of the instructional techniques included in the materials.

Subordinate competencies were postulated as follows:

> Ability to discriminate the use of such techniques in sample instructional sequences, both on a single and multiple discrimination basis.
>
> Ability to revise given sequences to improve the use of such techniques.
>
> Ability to correctly select statements relative to the recommended use of techniques.

1. Pretesting of criterion materials was conducted in February on two samples of subjects. One group was comprised by twenty-nine graduate students in curriculum at UCLA, the second consisted of eleven teacher education candidates at a local state college. Students were asked to discriminate the techniques in instructional sequences and also to define the terms presented. Pretest performance was found to be an experimental variation of the order in which the pretest tasks were presented, was conducted and reported subsequently in a paper (Baker and Quellmalz, 1977). In addition, a second experiment was conducted, comparing the strength of requiring criterion item-writing prior to instructional design, both experiments are reported in fuller detail in the section labeled "Controlled Variations."

2. The plan for the package of the materials was formulated:

> The prototype would consist of text materials, information items, practice exercises for discrimination, revision exercises, writing simulations and posttest materials. To differentiate materials for graduate students and to lend credence to the notion of research based instructional techniques, an extensive annotated bibliography was planned. Materials were intended to be rendered in print because of cost constraints.

3. A literature search was conducted for experiments relevant to instructional techniques selected for the project. Information from these experiments was abstracted and formed the basis for the text section. The notes were also to be refined for inclusion in the annotated bibliography at the end of the materials.

Phase Two: Tentative Text, Items, Revision and Draft Writing, April to [illegible]

1. Draft materials [illegible] advance organizers and control of inspection behavior texts. After editing, criterion items were produced which sampled the information presented in the texts. Practice exercises were either written or adapted to the topics.

2. During April, a developmental trial was conducted with the two components.

3. A first draft with appropriate text items and practice discrimination exercises was produced for the task description component.

4. An additional trial was conducted in June testing advance organizer, task description and control of inspection behavior texts and discrimination exercises. The materials were collated to form a subsection of the material entitled "Techniques for Directing Attention

5. During late May and early June, texts, criterion items and practice exercises were written for the second subsection, "Techniques for Response Control." The techniques of prompting, direct practice and knowledge of results formed the substance of the component.

6. Sets of specifications were produced for the simulation section of the materials. Directions for use of the specifications were also provided.

7. In late June, thirty copies of the materials were produced.

8. Work proceeded on the bibliography.

9. Contacts for field test sites were pursued.

Phase Three: Testing and Revision Cycles -- July-November 1971

1. Field test on Response Control section was conducted July 2 and 3 in Santa Monica.

2. Directions and items were revised and materials were tested at Arizona State on undergraduates. No staff members were directly involved in the field test.

3. Data from Santa Monica and Arizona were scored. Revisions were made in items, scoring procedures, practice discrimination exercises and text materials.

4. Field trial was conducted on July 27 and 29 in Torrance. Staff observed, but the training was supervised by local team leaders.

5. Data were compiled, analyzed and revisions in text, items, practice discrimination exercises were made.

6. The materials were subjected to external expert review by two educational psychologists. Their responses resulted in the simplification of language employed in materials as well as modification in the instructional sequence.

7. Revision exercises were produced.

8. Models of instructional sequences were developed and included in materials.

9. The pretest was modified to reflect revision rather than discrimination behaviors.

10. A field test was conducted at Manhattan Beach under optimal conditions. Data were considered to represent the first demonstration trial.

11. Data were analyzed and revisions made.

12. Pretest administered to UCLA teacher education candidates. Field trial not pursued.

13. Individual revision exercises (for each instructional technique) were produced.

14. Directions written for "hands-off" field test at Arizona State.

15. Modifications were made for laboratory field tests.

16. Revisions in directions for Pacific Bell Company field trial.

Phase Four: Final Data Analysis, Revision, and Report Preparation
December-March 1971

1. Lag in the receipt of data delayed production of final version.

2. The bibliography was criticized and seriously revised.

3. An optional section to organizers was reintroduced.

4. Introduction to materials was written which incorporated the history of development of this project.

5. The research studies generated through the project were compiled. A paper was prepared and presented at the Annual Meeting of the American Educational Research Association in Chicago.

6. Professional drawings were added to materials.

7. All data from all trials were subjected to reanalysis and graphing.

8. Based on data, minor organizational changes were instituted in materials.

9. Final copies of materials were generated, proofed and reproduced.

10. Preparation of final report began in March.

DEVELOPMENT OF CRITERION INSTRUMENTS

Text Criterion Items

Information items, which tapped immediate comprehension of the material presented in the text segments, underwent an extensive developmental sequence. The tension in the development plan derived from the need to have the items represent the critical areas of substance in the text while not requiring so many items that the field testing time would be enormously expanded. A prominent solution to such problems would invoke the use of item sampling techniques, where different items were distributed to different subjects. This solution was considered but eliminated both because of the clerical requirements it would impose on a thin staff and the restricted size of our field testing samples.

Items were generated in accordance with item form procedures (see Hively, Patterson, Page, 1968), in that the domains to be covered were identified and rules for generating incorrect responses were initially described. A standard item format was selected, requiring a multiple-choice response with four options. Each item was inspected following field trials. We had hoped to organize our field trials into pairs: the first trial was intended to provide data for the upgrading and refining test items; the second trial would provide information to be used for improving the instructional materials. Data for each item were arrayed on a card, with percentage of responses for each alternative displayed. Items with difficulties of less than .80 were flagged and carefully scrutinized. Possible ambiguities in the phrasing of certain items were identified. More often, however, revision for the text material itself was implied by the item data.

Placement of criterion items for text material was also varied. On early trials, items relevant to a major area, for instance, Techniques of Response Control, were presented together. On subsequent trials, however, the text criterion items were placed at the conclusion of the text materials to which they were relevant. Text criterion items should not be confused with the embedded information items which were inserted in the final version of the materials. The items interpolated in the text were to serve an instructional, rather than an evaluation function.

Information items for use on the pretest and on the retention test were randomly selected from the pool of text criterion questions which had been identified for all text sections.

Retention Test

Fifteen multiple-choice items, three each randomly selected from each of the questions relevant to the five text segments, comprised the retention test. The purpose of this test was to determine how much information was remembered at the conclusion of the entire training package. The completion of the text criterion items described above usually occurred within two or three hours of the onset of instruction. The retention test was administered at the close of the instructional sequence, from five to twenty hours after instruction commenced. Revision procedures for these items were identical to those described above, and especially poor performance (difficulty below .7), on the retention item was noted and revision of item or test was usually required.

Revision Criterion Task

Two sample instructional sequences were developed to serve as the bases for the revision task, where the learner would be required to demonstrate ability to correct instructional sequences with regard to deficiencies in the use of particular instructional techniques. Revision exercises were scored in three ways, one of which survived to the final trial. The first procedure required that the sequence be analyzed in terms of its peculiar deficiencies. For instance, if the technique of knowledge of results was missing, how many opportunities did the learner have to insert the use of the technique. The learner's score was a proportion of the number of times the technique was used over the number of opportunities available for use. Obviously, this procedure was appropriate for knowledge of results, but became more tenuous for the other instructional techniques. For example, a task description, if adequately stated, may need to be included only once. Since frequency rules for the use of text-embedded questions, practice and prompts were eschewed in favor of empirical trials, the arbitrary delineation of numbers in order to derive a proportion did not seem to be worth retaining.

A second procedure involved the assignment of effectiveness ratings from five to one by the scorers of the revision exercises. Effectiveness parameters were explicated for the scorers and revised so that agreement would hopefully be forthcoming. However, the necessary congruence among scorers on the assignment of effectiveness scores was never obtained. The last procedure employed, and ultimately retained was an on-off judgment of the use of each instructional technique. Agreement on the use of techniques was excellent, essentially perfect.

Writing Simulation

The writing simulation performance test was designed to measure adequately the objective of the project. While recognizing that the testing condition would be contrived in that it was both time and location controlled, the attempt was made to provide relatively realistic stimuli for the writing simulation task. The trainee thus was presented with a series of short-term instructional objectives suitable for the design of replicable instructional material. The learner was directed to employ, where possible, the instructional techniques advocated in the materials which they had completed. Sets of specifications were developed which included objectives for all grade ranges for a variety of subject areas. The criterion task would require the participant to select a subject field in his or her area of interest or expertise and to spend about one hour preparing a first draft of materials which exhibited the appropriate instructional techniques.

A major problem with a constructed response task as described above is deciding on an adequate scoring procedure. A series of scoring alternatives were developed and tried on the criterion data obtained from the field trials. An early version of the scoring sheet is presented in the Appendix. Rather complex decision rules were provided to guide the scorers' use of the guideline.

Responses from twenty-six subjects were scored by two scorers using the form above. Pearson product moment correlations were computed by technique.

TABLE 1

Correlation Between Two Scorers of Constructed Responses
By Instructional Techniques

Technique	r
Task Description	.84
Inspection Behavior	not used
Direct Practice	.49
Knowledge of Results	.74
Prompting	.28

Analysis of the scorers' problems through discussion revealed that the ambiguities occurred in assigning "effectiveness" points rather than in determining the presence or absence of the techniques themselves. The form, as generated, resulted in a subscore by technique which incorporated both "use" dimension and "effectiveness." A variant of this scoring procedure was adopted, where the scorer rated the presence or absence of each technique separately. In addition, an effectiveness rating, from 1 to 5, based on a set of explicit criteria was assigned to the use of each technique. While scores assigned by independent raters on the presence or absence of a technique were virtually identical, there was considerable variation in the correlations obtained between pairs of scorers. Because of the desire to have both the instructional and evaluation components of the materials approach standards of replicability, the final scoring scheme relied on the on-off detection of the presence of a given technique, with an optional effectiveness rating system. A copy of the form is included in the Appendix.

Attitude Measure

A measure of interest and attitude toward the materials was developed for use in formative evaluation. Questions regarding the length and complexity of the materials, clarity of explanation, and usefulness of ideas were designed to provide feedback to the development staff for future revisions. A copy of the program questionnaire is included in the Appendix.

In addition, a revision sheet was provided which consisted of four sections: (1) a space for the trainee to list specific suggestions regarding how the material should be revised; (2) a summary impression of the materials; (3) a summary statement of how interesting the materials were; (4) a personal evaluation of learning. Data from the revision sheet were used both as a basis for materials improvement as well as dependent measures for controlled variation work to be described in a subsequent section of this report.

INSTRUCTIONAL COMPONENT DEVELOPMENT

While the resume of work highlights the project activities chronologically, the explicit activities of the project need also to be described in more segmented terms. The following section will describe the tasks which were required for completion of the project and include elements of procedures which might prove useful to prospective instructional development personnel. Where relevant, examples of forms used are included in the Appendix.

Text

Once the topic for instruction had been selected, the development of the text segments proceeded along fairly conventional lines. The text sections were written in an attempt to convey accurate information regarding the utility and application of the various techniques. The purpose of each text section was to persuade as well as to inform the user regarding the value of employing the technique in instruction. The text length was limited by the number of discrete information points required by the given techniques, resulting in substantial length variations among techniques. For instance, the information statement for the technique of task description was considerably shorter and based on less theoretical complexity than the statement related to the control of inspection behavior. Among the rules to guide the production of the text section were the following:

1. Attempt to extrapolate accurately from referenced research.
2. Attempt to avoid oversimplification.
3. Avoid the use of citations and journal-style justifications in the text.

In addition, references to live teaching as well as instructional materials development projects were included, primarily to assuage the teacher education subjects employed in field tests.

During the major rewrite of the project (August, 1971), the language was simplified and each text section was reorganized to include the following categories:

1. Introduction: designed to motivate the reader. This section normally referred to prior school experiences of the trainee as a basis for understanding the value of the proposed technique.

2. Definition: the operational definition of the technique employed.

3. Forms of the technique: descriptions of alternative renderings of the technique.

4. Types of tasks: if possible, the variety of tasks for which the technique was particularly suited.

5. Effective use of technique: frequencies and contexts in which the technique optimally might be employed.

Special attention to the theoretical basis for the technique and to the types of learners for which the technique might be particularly appropriate were also included. During a special field trial, a comparison between the use of summaries and text-embedded adjunct questions was conducted and is described in detail in a later section. The final version of text materials, however, reflected both treatments: questions were inserted at relevant points in the text and brief summaries of the major points of the text material were added.

Practice Discrimination Exercises

Practice exercises referred to discrimination opportunities for the trainee where, within a given segment of an instructional sequence, the presence or absence of single or multiple use of the technique was to be determined by the learner.

The discrimination activity was assumed to be prerequisite to the ability to revise instructional sequences deficient in the use of given techniques and to the ability to generate instruction which included the specified techniques. An attempt was made to introduce variations into the instructional sequences used in the practice discrimination exercises, so that different aged learners, different types of tasks, and different instructional formats would be included. This decision was reached after a review was conducted on the initial practice exercises designed. These sequences depended heavily on cognitive classification tasks and were presented in a very linear "programmed instruction" format, which could easily promote mismodeling by the trainees.

Specifications to guide the development of the practice discrimination exercises were produced and revised. Samples of the single and multiple discrimination item forms are included in the Appendix.

As evidenced in the sample item forms, the original response required of the trainee was a simple yes or no discrimination regarding the use of the technique or techniques under scrutiny. It became apparent that learners might be able to respond satisfactorily to the exercise by merely skimming the sequence. The evil in such practice was suspected to lurk in the learner's reduced exposure to the range and nuance of use of the instructional techniques and the reduction in the modeling function these exercises might serve. To remedy the situation, learners were also asked to rate the "effectiveness" of the use of the technique, using for their criteria the recommendations from the text section. As knowledge of results sections were provided for the discrimination task, a quasi-knowledge of results for this rating behavior was devised. The project staff members were asked to rate the use of the technique. Their modal response provided the "expert" rating which was included in the knowledge of results for the practice exercises. Explanations for low ratings were provided as feedback to the trainees.

A major difficulty with the design of practice exercises was the inordinate time required of the learner to read through the sequences in order to respond to the exercises. Three exercises were designed for each instructional technique as well as two comprehensive practice exercises which combined the techniques of direct practice, knowledge of results and prompting and two exercises which combined task description and control of inspection behavior. The decision was made, when reading times and fatigue indicators soared on initial trials, to permit the trainee to choose to view these exercises as remedial. Another option was that the trainee complete two of the three exercises for each technique, and if correct on both, skip the third exercise in the set.

Instructional Models

In our early formulation plan the practice discrimination sequences were to serve multiple purposes. First, they were to provide opportunity for the participant to identify the presence or absence of a particular technique in a simulated instructional sequence. In the case of positive instances, they were also to serve a model function, by demonstrating what the use of the technique would "look like." However, during the course of the tryouts, suggestions from subjects often raised the need for a model or sample of reasonable use of the techniques. To this end, a set of models was prepared and inserted following the practice discrimination exercises. These models were to employ the techniques in linear situations, so that the learners could refer to them as concrete examples of the abstractions in the set. An added feature of the models was a written supplement, which would explain what was happening in the sequence, e.g., "the technique of practice was employed when the problem situation was posed." These explanations appeared juxtaposed to the instructional sequence at the precise location where the use of the technique occurred. Reference to the models was designed to be optional and appropriate for those learners who needed more examples than those furnished by the practice discrimination exercises.

Revision Exercises

The ability to locate and remediate specific deficiencies in instruction was thought to be prerequisite to the behaviors required in the writing of first draft instruction which includes the use of particular instructional techniques. For our initial revision exercises, two instructional sequences were produced. Each included incorrect use or sometimes the absence of use of all five instructional techniques. Learners were asked to review each sequence and its specifications and to improve the material according to the recommendations provided in the text sections. A series of difficulties in the design and implementation of these exercises was uncovered. First, when the provided instructional sequence was poorly written across the board, or when it represented a particularly trite approach to the teaching of the skill, the trainees attempted to rewrite the entire exercise. The time they spent in the reformulation of format interferred with their correction of the instruction along the guidelines expected by the exercise. A second problem related to the nature of the task. When directed to correct or improve a sequence according to given instructional techniques, some of the participants tended to write notes, such as "I'd add more practice" or "I'd tell them what the objective was" without demonstrating that they could do an adequate job of the task. While the identification of a deficiency is clearly prerequisite to the design of a remedy, it is not sufficient. The directions provided the trainees were improved and structured so that identifying locations of deficiencies and the specific remedies proposed were required of the respondent.

Suggestions from participants following the field tests indicated that the transition from multiple discrimination to multiple correction was too abrupt. The sequence was modified so that five additional revision exercises were inserted following the multiple discrimination tasks and prior to the comprehensive revision exercises. One exercise for each instructional technique was produced, so that the learner needed to focus only on the use of task description, for example, in his or her initial attempt at correcting deficiencies. These sequences exhibited a stronger series of prompts, in that the learner's attention was directed only to a single technique and that it was clear that there were errors in its application in the instructional sequence. The comprehensive revision exercises required that the learner synthesize two skills: first, multiple discrimination, to determine which, if any, techniques were employed inappropriately or not at all; and second, to make the correct applications of the technique.

Confirmation of these exercises was provided by a series of alternative models, regarding what modifications were deemed appropriate by the staff. Paper exchange with a peer was also encouraged.

Writing Simulations

The development of the writing simulations closely adhered to the procedures used in preparing the criterion testing situation on first-draft production. Sets of specifications were produced which exhibited content appropriate to various subject matters and grade levels.

An important dimension of the simulation exercise was providing a range of specifications from which the learners could choose. We did not wish to fail in our task because we had identified only a narrow band of subject matters and age levels for the first-draft writing practice. We also wished to avoid the freest situation, where the learner would generate the specifications for his or her own work. Previous training experience indicated that the specifications which were generated by the learner would suffer with high probability from a number of deficiencies. The specifications might not be operational or stated in sufficient detail. They might be such that they required only simple responses from the intended learners and thus not provide a task of sufficient complexity so that the instructional techniques would have the opportunity to be displayed. Short of attempting to get the trainees' specification-writing behavior under the program's control, a task which would deplete our staff's resources, the provision of a wide range of alternative specifications from which the trainee could select his practice situations was accepted as a reasonably appropriate stimulus condition. Further rationalization for the process was derived from previous product development work, where the writer of instruction often has only minimal input into the design of instructional specifications and more often has these specifications provided by the management level of the project.

A number of sets of directions was written and tried out with successive field test groups. In addition, the amount of time allocated for the practice in product writing was varied. In early trials, only a single half hour was permitted, usually because of time constraints. In later versions, the amount of time was expanded to over an hour. The final trials permitted 45 minutes to one hour for each product writing and encouraged two attempts by the participants.

The confirmation of such practice opportunities posed a few problems. Obviously, since the responses were to be constructed by the participants, some notion of the adequacy of their attempts needed to be communicated. The provision of a range of models was one solution which was considered. Following the writing, a series of sample sequences relating to each specification would be presented. The participant would compare the work he or she produced with the models presented and determine whether the techniques were incorporated in them. This solution was rejected on cost-effectiveness grounds. On the one hand, it was feared that the provision of samples would unnecessarily limit the trainees' view of proper use of the instructional techniques; on the other hand, the amount of staff time required to develop adequate samples for each of the specifications which might be selected would place excessive demands on our capability. Instead the procedure of peer confirmation was employed. Trainees were instructed to work in pairs with another participant. Following the designated writing period, a ten-minute exchange and critique period was provided, where the trainees would read each other's efforts and criticize the use of the techniques. Where questions arose, the trainees referred to the text materials provided. Observation of this peer confirmation practice revealed no major difficulties. People were evidently able to exchange and criticize each other's work profitably.

Annotated Bibliography

The development of the annotated bibliography was undertaken because it was believed that such a component would contribute significantly to the differentiation of the product's utility for various groups. Graduate students, for instance, would be especially interested in pursuing some of the allegations made in the text sections regarding the various techniques, although it appeared unlikely that school personnel would be interested, in general, in the research design, sampling techniques and detail of experimental procedure. References relating to the given techniques were collected by research assistants. The studies were then reviewed to eliminate apparent overlap or irrelevant experiments. Abstracts of the research studies were prepared which included the following categories: problem or hypothesis, subjects, sampling procedure, task, treatment length, results, and conclusions. Although the actual abstracts provided the information with certain variability in description, the overall purpose was to whet the reader's appetite for more such experimental writing. The component, because it was not essential to the accomplishment of the project objectives, was deferred and was completed later in the project's lifespan. While not formally tested, the bibliography has undergone two careful reviews and a total rewrite. On the basis of local interest and demand at UCLA, the bibliography is anticipated to be a valuable addition to programs on instructional research and development.

The total package including all the components required approximately fifteen hours of instruction and testing to complete.

TABLE 1

Distribution of Time by Instructional Component

Component	Minutes (approximate)
Text, including information items	[illegible]
Practice Discrimination Exercises	[illegible]
Models	[illegible]
Revision Exercises	
Simple Technique	[illegible]
Comprehensive	[illegible]
Writing Simulation	[illegible]
Posttest	[illegible]
Pretest	[illegible]
TOTAL	[illegible] hours

PRODUCT TRYOUT

The purposes of the product tryouts varied with the point in the development sequence where they occurred. Early trials were to verify the suitability of prototype formats and criterion items. Later trials served to identify gaps in the sequence and to provide data for revision of previously developed materials. The trials also were designed so that progressively less staff involvement in the tryout was required, in order to ascertain whether any degree of program replicability had been achieved.

Subjects

Not totally by design but partly by happenstance, the subject population varied by trial. The staff believed that the materials should be tested on different kinds of people with various job requirements which pertained to instructional development. Subjects participating in the field trials represented four distinct groups: preservice teacher education candidates, graduate students in instructional development, experienced school personnel, and instructional development staff performing in outside agencies. Colleges and universities provided the source of students at UCLA (N=29) and teacher education candidates at San Fernando Valley State College (N=11). The first prototype trials of the directive attention component was tried on two groups of UCLA teacher education candidates (N=10; N=8). The prototype materials on response control were tested on experienced elementary level school teachers (N=). The first off-site trial was conducted at the Arizona State University campus on undergraduate education students (N=18). Curriculum supervisors on the central staff of a medium-sized school district tested revised components of the materials (N=). The internal validation trial was conducted in cooperation with the Manhattan Beach School District and California State College, Dominguez Hills, employing interns in teacher education (N=11). Arizona State College graduate students served as one external validation sample for the alternational version of the materials (N=11). The final sets of data were obtained from a trial on staff course-writers employed by the Pacific Bell Telephone Company, Training Division (N=14). These individuals deviated markedly from the other groups in terms of educational experience and job requirements

Subject Procurement

Attempts to obtain subjects occupied a substantial portion of staff attention throughout the grant period. A flyer was developed which attempted to solicit participants from various school districts (see Appendix). Such requests were often transferred from person to person within the structure of a school district and rarely resulted in a field test site. Among the salient injuries sustained in this regard was an attempt to use UCLA teacher education students in a validation trial October and November, 1971. Their education courses were being conducted on a highly individualized basis. The students rebelled when asked to complete a pretest in a group situation and anxiety and negative affect on the part of the potential subjects and the staff members were at a level where it was felt most productive to abort the field test.

A second area of disappointment was the inability to secure participation from regional laboratory staff. Requests were sent to six regional laboratories and positive response were received from four. Packets and directions were prepared and mailed, but no further word was received from three of the sites. The fourth, the Far West Regional Laboratory, had intended that the materials be tried on participants in a joint project with the American Institutes for Research, however, limited instruction precluded the use of the materials and the packets were returned unopened. Explanations for these responses are many; perhaps the materials were seen to be irrelevant; perhaps the mission orientation of the development agency interferred with the ability to free staff for the approximately fifteen hours required for program training. Plans are underway to query the regional laboratory staff members to determine the reasons for the lack of completion of the materials.

A final point should be made regarding the acquisition of subjects and field testing sites. In all but one of the field tests conducted, the contact was initiated by the project director or her coordinator and directed to someone with whom previous successful contact had been made. The percentage of positive results of contacts with new people in laboratories, school districts, and other institutions was depressingly low.

Testing Schedule

The program was tested in components during the various product tryouts. The following table presents the date, location, number of subjects and components tested at each tryout.

TABLE 3

Testing Schedule

Date	Location	N	Components Tested
February	UCLA	29	Pretest
February	San Fernando Valley State College	11	Pretest
April	UCLA	10	Advance organizer, inspection behavior text and discrimination exercises
June	UCLA	18	Entire component of directing attention including texts, items, and discrimination exercises
July	Santa Monica	9	Response control text, items and discrimination exercises and writing simulation prototype
July	Arizona State University	18	Entire package to date: including text items and discrimination exercises for both response control and directing attention, writing simulations
July	Torrance	7	Response control package plus revision exercises
September	Manhattan Beach	21	Entire package, including texts, items discrimination exercises, models, revision exercises and writing simulations
November	Arizona State University	22	Same components as above
December	Multiple California Locations	14	Entire package

In addition, a controlled variation test was made during the Winter, 1972 on the use of embedded questions (N=22). The text materials and items were the only portions of the program employed. A supplementary test of part of the materials was made during January during an experiment comparing media variations of the discrimination exercises (N=50). This study will be reported in a separate section.

REVISION ACTIVITY

The bases for revision were many. They included data from the tryouts of the program on appropriate subjects, reviews by individuals external to the project staff and judgments made by the staff members. Because ten separate tryouts had been planned (although only eight materialized) the revision pattern was intended to be controlled by the sequential acquisition of data. Ideally, each version of the program would be revised as a consequence of judgments made based on data acquired from the most recent field trial. Thus, data from Trial One should contribute to the development of Version Two and so on. Unfortunately, procedural problems conspired to upset such progressive application of performance data. Field trials, for instance, needed to be scheduled well in advance, so organizations could commit subjects for the length of time estimated. Therefore, the development team was always working against a schedule: the dates that subjects would be ready to begin the program. Second, many of the trials were planned off-site, so that the staff was dependent upon the program administrator to mail the raw data back to them. Often, there were inopportune delays. The multiple responsibilities of the staff members also compounded the problem, for the project could not afford a single person with only a data analysis task. Data analysis responsibilities may have been perceived as interrupting more continuous and interesting development or "creative" activities.

Thus, the inevitable data analysis lag from the previous tryout, when combined with the preordained tryout schedule and the enormous time required to prepare multiple copies of materials resulted in a situation where data from one field trial were being tabulated at the same time the materials for the next field trial were being prepared. Rather than on the next proximate field test, data-based revisions tended to be incorporated in materials at a later time. Thus, data from Trial Four had its impact on Trial Six.

A revision strategy was employed where we could take account of information from the most recent tryouts without unduly deferring material preparation and multiple copy production. Data from the immediately preceding tryout were employed in two ways. First, criterion items were revised. Item editing and revisions in scoring procedures were possible without delaying costly production efforts. Instructions to the user, usually contained in brief letter format, were also revised, as were directions given to the program administrator at the field test site.

The major revision activity occurred in August and September, 1971. During this period, a rewrite of the entire set of text materials was completed, based, in part, on the accumulated data from Trials One through Five and the detailed review submitted by Dr. Howard Sullivan of Arizona State University. Subsequent modifications to the project from Trials Six through Eight included variation in sequence and directions to the user, addition of specifications for the simulation tasks, inclusion of the bibliography and technical modifications in visual design.

The summary of data-based revisions is presented below.

Version One April, 1970

Text Instruction:

1. Final selection of instructional techniques.

2. Generation of Item Forms.

3. Preparation of expository instruction for selective attention component.

Criterion Items:

Preparation of information items.

Practice Exercises:

1. Identification of practice exercises suitable for selective attention.

2. Adaption of practice exercises to:

   a. Reflect the entire scope of the objective.

   b. Employ positive or negative instances of selective attention techniques.

   c. Stylistic improvements.

Version Two June, 1970

Text Instruction:

1. Preparation of instruction on Task Description.

2. Clarification of Directing Inspection Behavior and Advance Organizer as indicated by trial data on criterion items and practice exercises.

3. Stylistic editing.

Criterion Items:

1. Revision of items as indicated by empirical data.

2. Addition of items for Directing Inspection Behavior.

Practice Exercises:

1. Graphic cues and prompts added to Practice Exercises on which low performance occurred.

2. Directions for responding to Practice Exercises clarified.

3. Elimination of nonessential instruction in Practice Exercises.

Version Three July, 1970

Text Instruction:

1. Revision of all instruction in Techniques for Directing Attention to eliminate the majority of theoretical rationales and technical editing.

2. Stylistic editing.

3. Shortening of Control of Inspection Behavior as indicated from program questionnaire data.

4. Preparation of text instruction for Component II--response control, including practice, prompting, knowledge of results.

Criterion Items:

1. Revision of Directing Attention items as indicated by data.

2. Preparation of additional items for Advance Organizer and Task Description.

3. Preparation of items for Response Control instructional segments.

Practice Exercises:

1. Replacement of three Techniques for Directing Attention practice exercises whose programmed format and/or subject matter complexity rendered them inappropriate for attention direction techniques.

2. Revision of remaining practice exercises as indicated by trial data.

3. Preparation of Practice Exercises for Techniques for Response Control.

Simulations:

Elementary level specifications for writing simulations generated.

Version Four July, 1970

Text instruction:

No change.

Criterion Text Items:

1. Posttest (retention) constructed.
2. Revised for Response Control segment.

Practice Exercises:

No change.

Simulations:

Number of specifications expanded.

User Directions:

Orientation for leader (nonproject staff member) of session prepared.

Version Five September, 1970

Text Instruction:

1. Complete rewrite, simplifying language of text materials.
2. Advance Organizer deleted.
3. Revisions where test performance below .80.

Text Criterion Items:

Rewritten.

Practice Exercises:

Edited.

Simulations:

Specifications expanded.

Revision Practice Sequences:

Added on Posttest.

Model Instructional Sequence Prepared:

1. Directions for supervisors revised.

2. User letter prepared.

Version Six October, 1971

1. Materials resequenced to be ordered by technique, e.g., direct practice, rather than mode, e.g., text.

2. User letter revised
   Supervisor instruction revised

3. Practice exercises rewritten where potential copyright infringement existed.

4. Revision exercises relevant to each technique added.

5. Writing Simulations:

   Secondary level specifications added to practice and posttest simulations.

Version Seven Final Version January, 1972

Introduction rewritten

"Organizer" section revised and added as an option

Bibliography added

Directions to user revised

DATA ANALYSIS AND RESULTS

Data from each tryout are presented in sequential tables below. Tables reflect performance on instructional exercises as well as criterion tasks. The affective data are reported in two summaries with percentages for both positive and negative items presented.

In reviewing the progress of the project, it is well to keep in mind that the subjects varied greatly in experience and predisposition from tryout to tryout. Secondly, each field test generally reflects an increasing time and content load on the participants. While performance cannot be directly compared from trial to trial, because of test item revisions as well as sample variation, a sense of the progress of the project can be obtained by following the empircal history of the project.

Technical features of the tables include the use of percentages as the most comprehensible transformation, the summary by technique for both instruction and criterion, to determine the differential difficulty subjects experienced with the material. The number of subjects for summary data varies from that described as the number participating in the tryout because only information from those individuals with no missing data were used in the analysis. Table 4 presents pretest data while the following tables display results from tryouts.

TABLE 4

Pretest Data

	Combined Across Two Groups		
	$\bar{X}$	%	N
Discrimination *	13.7	55	40
Information **	1.3	13	40
Writing ***	4.6	26	29

Total Possible:
* 25
** 10
*** 18

Simple graphs of the percentages by component across trials are presented in the following pages in an attempt to make the results more understandable.

TABLE 5

Trial One: Performance on Text Information Items and Practice Discrimination Exercises by Technique

Task	No. of Items	$\bar{X}$	%	N
Text Items:				
Task Description	7	4.31	57.1	10
Control of Inspection Behavior	5	3.79	77.8	10
Advance Organizer	5	3.00	65.3	10
Practice Discrimination Exercises:				
Task Description	3	1.39	45.5	10
Control of Inspection Behavior	3	1.26	42.1	10
Advance Organizer	2	1.57	78.9	10

TABLE 6

Summary Performance Data: Trial One

Component	No. of Items	$\bar{X}$	S.D.	$\bar{\%}$	N
Total Text Items	17	11.1	2.66	65.3	10
Total Practice Discrimination Exercises	8	4.21	1.74	52.6	10
Affective Data:					
Positive items	4	3.20	.27	64.0	10
Negative items	7	2.73	.54	54.6	10
Total by Technique:					
Task Description	12	6.89	2.18	57.4	10
Control of Inspection Behavior	10	6.21	1.78	62.1	10
Advance Organizer	12	5.52	1.80	61.4	10

TABLE 7

Trial Two: Performance on Text Information Items and Practice Discrimination Exercises by Technique

Task	No. of Items	$\bar{X}$	S.D.	%	N
Text Items:					
Task Description	7	5.39	.78	77.5	18
Control of Inspection Behavior	10	6.94	1.30	69.4	18
Advance Organizer	5	3.78	.65	75.4	18
Practice Discrimination Exercises:					
Task Description	3	1.94	.78	64.7	18
Control of Inspection Behavior	3	1.94	.78	64.7	18
Advance Organizer	3	2.22	.808	75.3	18

TABLE 8

Summary Performance Data: Trial Two

Component	No. of Items	$\bar{X}$	s.d.	%	N
Total Test Items (excluding organizer)	17	12.33	1.24	72.5	18
Total Practice Discrimination Exercise*	6	3.89	1.5	64.8	18
Affective Data:					
Positive Items	7	2.51	.278	70.7	18
Negative Items	4	1.14	.753	62.8	18
Total by Technique:					
Task Description	12	8.11	1.00	68.0	18
Control of Inspection Behavior	15	10.17	1.76	68.0	18
Advance Organizer	10	7.33	1.41	73.3	18

TABLE

Trial Three: Performance on Text Information Items, and Practice Discrimination Exercises by Technique

Task	No. of Items	$\bar{X}$	S.D.	%	N
Text Items:					
Direct Practice	5	4.0	.96	80.0	9
Knowledge of Results	5	3.0	1.33	60.0	9
Prompting	5	4.33	.67	86.7	9
Practice Discrimination Exercises:					
Direct Practice	3	3.00	.00	100.0	7
Knowledge of Results	3	1.57	.49	78.5	7
Prompting	3	3.00	.00	100.0	7

TABLE 10

Trial Three: Performance Information Posttest by Technique

Delayed Posttest:	No. of Items	$\bar{X}$	S.D.	%	N
Direct Practice	3	2.67	.52	88.67	6
Knowledge of Results	3	2.0	.63	66.70	6
Prompting	3	1.83	.98	61.10	6
Task Description	3	2.17	.75	72.20	6
Control of Inspection Behavior	3	1.83	.98	61.10	6
Advance Organizer	3	1.00	.63	33.33	6

TABLE 11

Summary Performance Data: Trial Three

Component	No. of Items	$\bar{X}$	S.D.	%	N
Total Text Items	15	11.33	2.49	75.53	9
Total Practice Discrimination Exercises	6	5.57	.49	92.83	7
Affective Data:					
Positive Items	4	3.72	1.04	74.4	9
Negative Items	7	2.37	.98	47.3	9
Delayed Posttest	15	10.50	2.36	70.0	6
Total by Technique:					
Direct Practice	7	6.00	.93	85.71	7
Knowledge of Results	7	4.86	1.12	69.43	7
Prompting	7	6.43	.73	91.86	7

TABLE 14

Summary Performance Data: Trial Four

Components	No. of Items	$\bar{X}$	S.D.	%	N
Total Text Items	32	22.14	4.09	69.19	14
Total Practice Discrimination Exercises	10	8.43	1.29	84.3	6
Affective Data:					
Positive items	4	2.41	1.48	48.10	8
Negative items	7	3.13	1.63	62.6	5
Total: Posttest	15	9.27	2.67	61.77	15
Total by Technique:					
Direct Practice	7	6.46	.63	92.28	13
Knowledge of Results	7	5.00	.91	71.43	12
Prompting	7	5.67	1.08	81.00	12
Task Description	9	6.43	1.59	71.44	7
Control of Inspection Behavior	9	8.33	1.89	92.56	6
Advance Organizer	9	4.33	1.38	48.11	6

TABLE 15

Trial Five: Performance on Text Information Items and Practice Discrimination Exercises by Technique

Task	No. of Items	$\bar{X}$	S.D.	%	N
Text Items:					
Direct Practice	5	4.88	.35	97.6	8
Knowledge of Results	5	3.25	1.28	65.0	8
Prompting	5	3.5	.93	75.0	8
Task Description	7	5.00	1.58	71.43	5
Control of Inspection Behavior	10	7.00	1.41	70.0	5
Advance Organizer	5	1.8	1.92	36.0	5
Practice Discrimination Exercises:					
Direct Practice	3	3.00	0.00	100.0	7
Knowledge of Results	3	2.86	.39	95.33	7
Prompting	3	2.71	.49	90.33	7
Task Description	3	2.33	1.15	77.67	3
Control of Inspection Behavior	3	2.00	1.00	66.67	3
Advance Organizer	3	2.00	1.00	66.67	3

TABLE 16

Trial Five: Performance On Information of Delayed Posttest by Technique

Delayed Posttest	No. of Items	$\bar{X}$	S.D.	%	N
Direct Practice	3	2.83	.37	94.3	6
Knowledge of Results	3	1.50	.96	50.0	6
Prompting	3	1.67	.47	55.6	6
Task Description	3	2.50	.50	83.3	6
Control of Inspection Behavior	3	2.33	.94	77.67	6
Advance Organizer	3	1.83	.69	61.00	6

TABLE 17

Summary Performance Data: Trial Five

Components	No. of Items	$\bar{X}$	S.D.	%	N
Total Text Items	32	23.80	2.79	74.38	5
Total Practice Discrimination Exercises	15	12.67	1.89	80.47	3
Affective Data:					
Positive items	4	3.19	1.18	63.74	4
Negative items	4	2.86	1.38	57.20	4
Total: Posttest	15	10.83	2.03	72.20	6
Total by Technique:					
Direct Practice	8	..79	.20	100.0	7
Knowledge of Results	8	6.14	1.24	76.75	7
Prompting	8	6.14	.99	76.75	7
Task Description	10	6.67	1.70	66.70	3
Control of Inspection Behavior	13	8.33	2.05	64.70	3

TABLE 18

Trial Six: Performance on Text Information Items and Practice Discrimination Exercises by Technique

Task	No. of Items	$\overline{X}$	S.D.	%	N
Text Items:					
Direct Practice	10	8.05	1.25	80.50	21
Knowledge of Results	5	4.58	.59	91.60	19
Prompting	6	4.89	.66	81.50	18
Task Description	7	6.00	.97	85.71	19
Control of Inspection Behavior	10	8.33	1.45	83.33	18
Practice Discrimination Exercises:					
Direct Practice	3	.95	.12	95.28	14
Knowledge of Results	3	.98	.09	97.50	12
Prompting	3	.92	.14	92.40	13
Comprehensive: Response Control	6	.90	.14	90.40	14
Task Description	3	.90	.15	89.80	13
Control of Inspection Behavior	3	.88	.21	88.00	11
Comprehensive: Directing Attention	2	.96	.14	95.80	12

TABLE 19

Trial Six: Performance on Information Delayed Post-test and Writing Simulation by Technique

Delayed Posttest	No. of Items	$\bar{X}$	S.D.	%	N
Text Items:					
Direct Practice	3	2.05	.67	68.3	21
Knowledge of Results	3	2.29	.78	76.3	21
Prompting	3	2.67	.58	89.0	21
Task Description	3	2.29	.78	76.3	21
Control of Inspection Behavior	2	1.29	.46	65.50	21
Writing Simulations:					
Direct Practice	1	1.00	0.00	100.0	19
Knowledge of Results	1	.95	.22	95.0	20
Prompting	1	.65	.48	65.0	20
Task Description	1	.84	.36	84.2	19

TABLE 20

Summary Performance Data: Trial Six

Components	No. of Items	$\bar{X}$	S.D.	%	N
Total Text Items	38	32.25	3.13	84.87	16
Total Practice Discrimination Exercises	23	.93	.08	93.10	14
Total Information Posttest	14	10.57	2.18	75.50	21
Total Writing Simulations	4	.89	.13	89.10	20

TABLE 25

Trial Seven: Percentage by Technique for Instructional and Criterion Components

Components	% across text components	% across post-text components
Total by Technique		
Direct Practice	85.65	85.65
Knowledge of Results	75.63	82.72
Prompting	89.88	72.60
Task Description	81.05	72.70
Control of Inspection Behavior	84.80	83.50

TABLE 26

Trial Eight: Performance Text Information Items and Practice Discrimination Exercises by Technique

Task	No. of Items	$\bar{X}$	S.D.	%	N
Text Items:					
Direct Practice	10	8.14	1.56	81.4	14
Knowledge of Results	5	3.43	1.16	68.6	14
Prompting	6	5.57	.76	92.8	14
Task Description	7	5.50	.94	78.5	14
Control of Inspection Behavior	10	7.29	1.64	72.9	14
Practice Discrimination Exercises:					
Direct Practice	3	1.00	.00	100.0	14
Knowledge of Results	3	1.00	.00	100.0	14
Prompting	3	1.00	.00	100.0	14
Comprehensive: Response Control	2	.95	.12	95.30	14
Task Description	3	1.00	.00	100.0	14
Control of Inspection Behavior	3	1.00	.00	100.0	14
Comprehensive: Directing Attention	2	.96	.13	96.40	14

TABLE 27

Trial Eight: Performance on Information Delayed Post-test and Writing Simulations by Technique

Delayed Posttest	No. of Items	$\bar{X}$	S.D.	%	N
Text Items:					
Direct Practice	3	2.07	.92	69.0	14
Knowledge of Results	3	2.14	.95	71.3	14
Prompting	3	2.64	.50	86.0	14
Task Description	3	2.29	.61	76.3	14
Control of Inspection Behavior	3	1.43	.65	71.5	14
Writing Simulations:					
Direct Practice	1	1.00	.00	100.0	8
Knowledge of Results	1	.75	.43	75.00	8
Prompting	1	.13	.33	12.50	8
Task Description	1	.75	.43	75.00	8

TABLE 28

Summary Performance Data: Trial Eight

Components	No. of Items	$\bar{X}$	S.D.	%	N
Total Text Items	38	29.93	4.32	78.76	14
Total Practice Discrimination Exercises	19	..98	.04	98.00	14
Affective Data:					
Positive items	4	2.43	1.45	48.60	13
Negative items	7	2.80	1.46	56.00	5
Total Posttest:					
Text Items	14	10.57	2.50	75.50	14
Simulations	1	.73	.14	72.50	8

TABLE 29

Trial Eight: Percentage by Technique for Instructional and Criterion Components

Components	% across text components	% across post-test components
Total by Technique		
Direct Practice	90.70	84.50
Knowledge of Results	84.30	73.15
Prompting	96.40	49.25
Task Description	89.25	75.65
Control of Inspection Behavior	86.45	71.50

FIGURE 1

Text Information Items

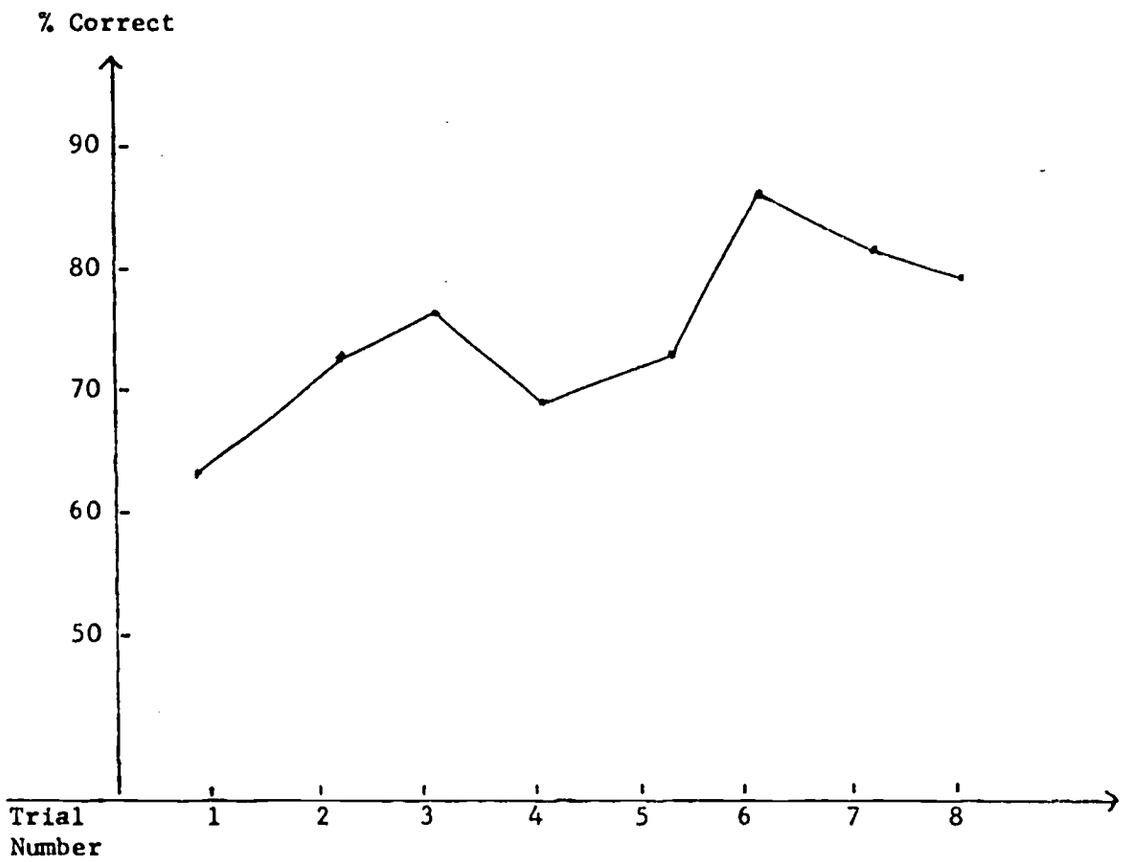


FIGURE 2

Practice Discrimination Exercises

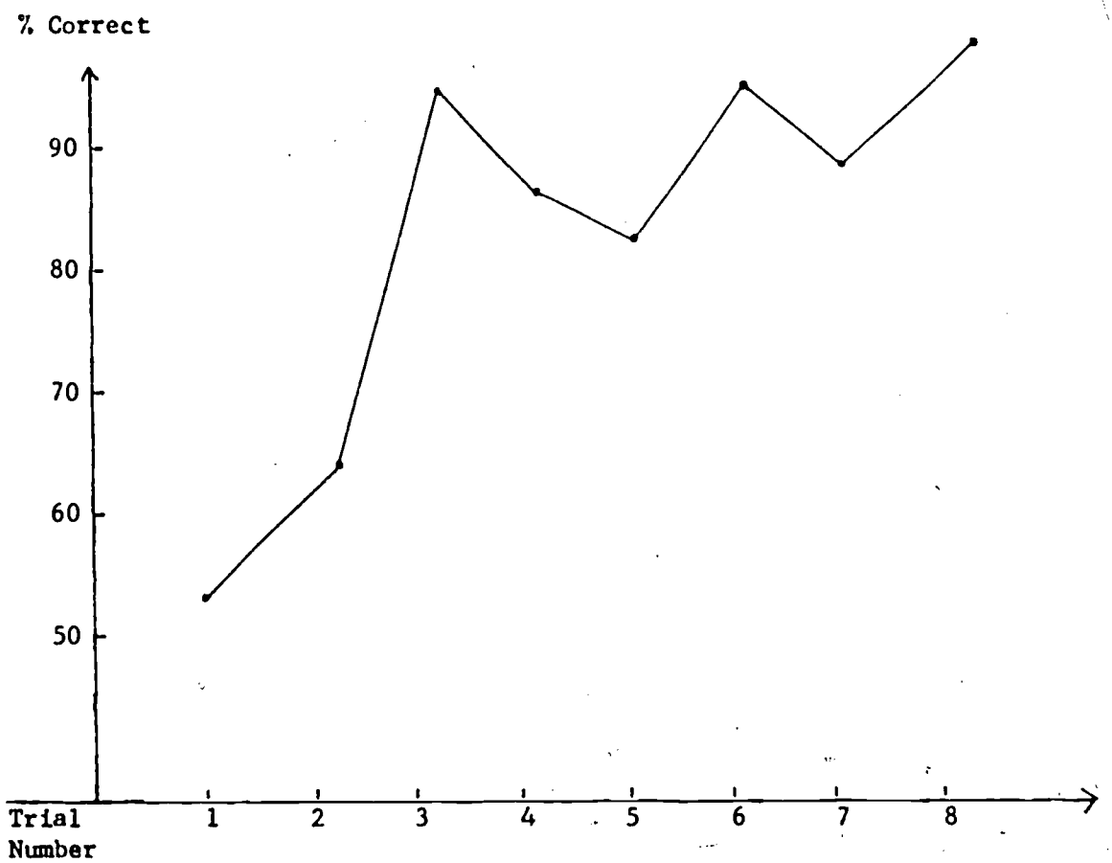


FIGURE 3

Delayed Information Test

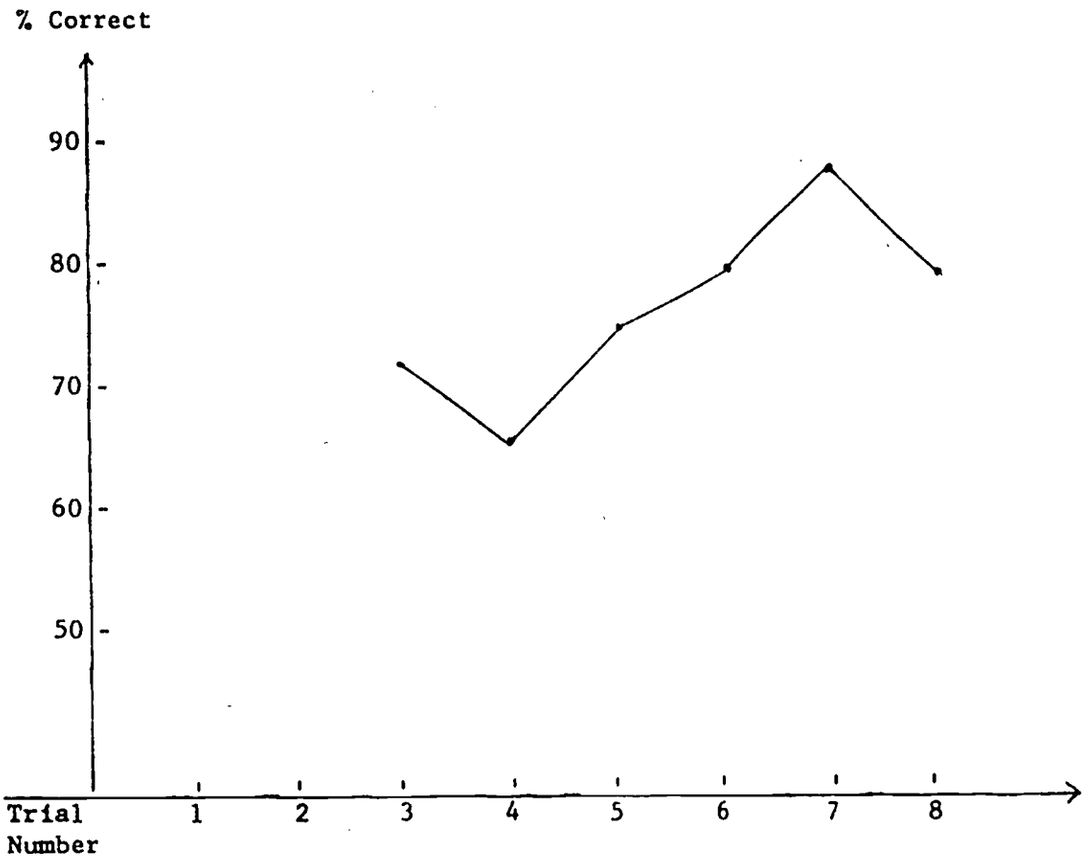


FIGURE 4

Writing Simulations

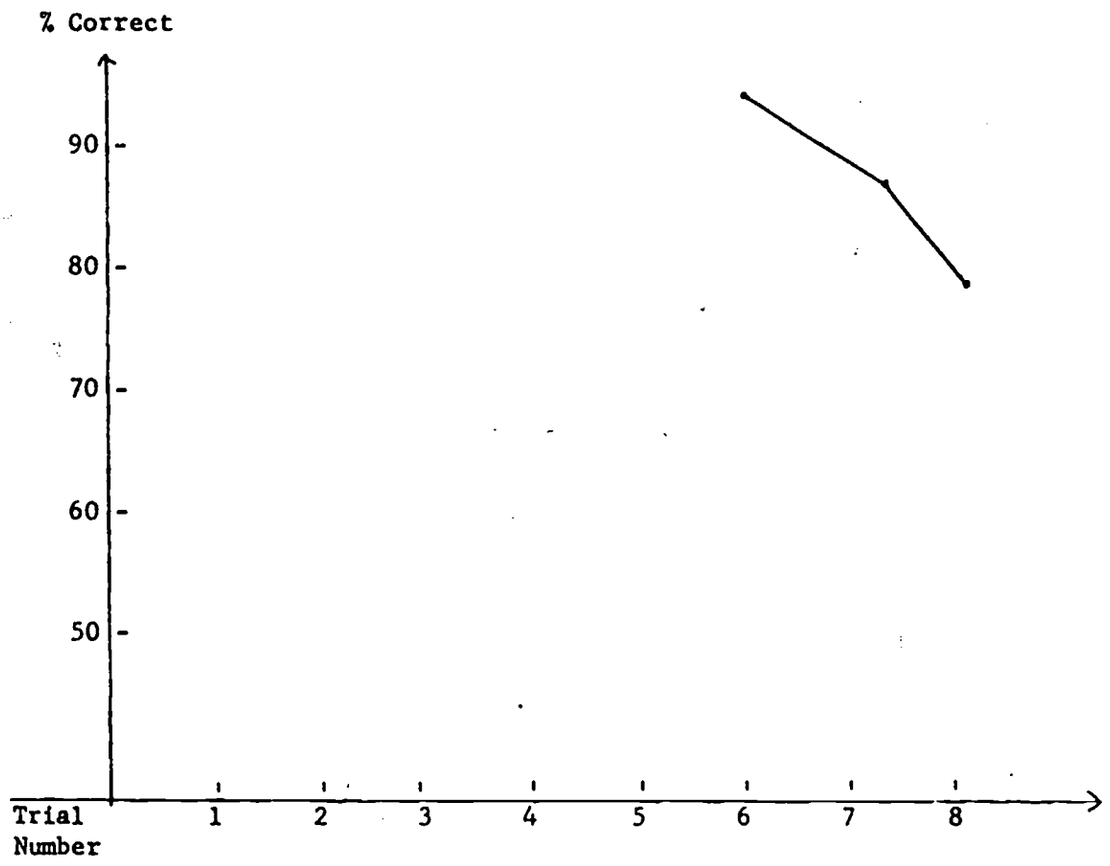
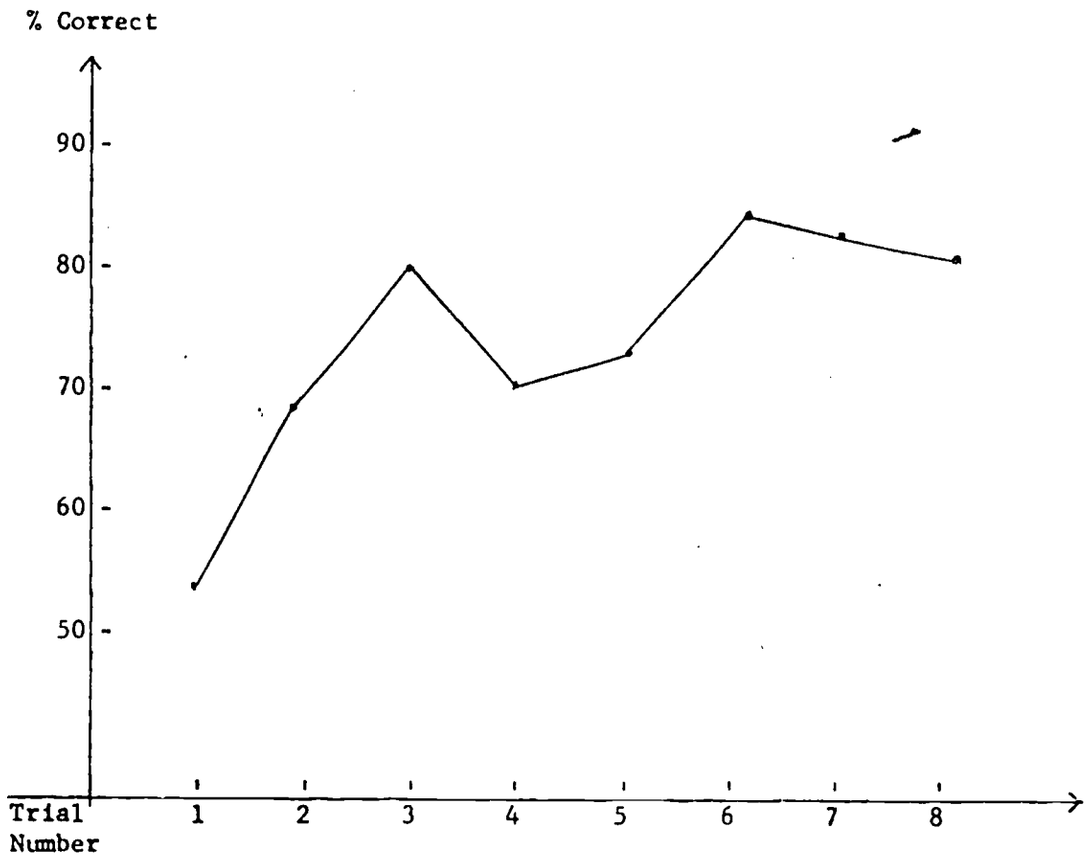


FIGURE 5

Summary Percentages of Materials Across Techniques and Objectives



Summary of Results

By inspecting the tables one may discern that in trials six and seven, the "validation" trials, most components are working well. Performance on the criterion simulation task was found to be more than satisfactory. Scores on information items, both on immediate and delayed basis were strong. Practice exercise performance was excellent. Performance on the revision task was less positive, but in light of the good showing on the simulation task, perhaps revision expertise is not prerequisite to the ability to employ the techniques. Attitude measures showed some improvement over early trials. However, since the materials would be used generally by highly motivated individuals, such as graduate students or those seeking to learn job-related skills, the lack of unabashed enthusiasm is tolerable. It might be noted that all field tests were conducted using dittoed copies of material, and none of the groups were exposed to professionally "finished" copy. Thus, it would be expected that the visual design improvements would contribute to the affective scores on subsequent trials.

Among the techniques, some consistent differences were observed. The technique of prompting never appeared to be strong in empirical tests regarding its application. In reviewing the text, however, we noted that readers are encouraged to avoid the use of prompts when they prepare early versions of instruction, and are directed to augment sections which demand strong stimulus control after empirical tryout. Since firstdrafts of instruction were requested, the poor performance with regard to this single technique was builtin as a consequence of the text instruction we provided. We elected not to change the instruction in order to inflate our criterion scores. The use of embedded questions to control inspection behavior also appeared less consistently than the other techniques. This fact was accounted for by the format in which the subject chose to cast his or her instruction. Not many of the subjects elected to use a standard text approach to the preparation of first draft.

We also maintained records of reading times for the materials, and noted that while the number of pages of materials tended to increase, the reading times did not increase similarly. In fact, after the major revision effort, preceding the Trials Six, Seven and Eight, reading times decreased. Evidently, the language simplification and reorganization of material was useful in making learning from the prose less time-consuming.

CONTROLLED RESEARCH WITHIN THE DEVELOPMENT PROCESS

In the course of the project's development, a series of research variations were embedded within the continuing field trials that composed the project's trial/revision process. Generally, the intent of the studies fell into one of two categories: (1) to clarify conditions under which learning techniques treated in the text materials would operate in applied instructional settings, (2) to explore empirically alternative product development procedures.

In some cases results were inconclusive, attributable in part to small samples of subjects available for developmental testing, or procedural aberrations. However, the contribution of the successful studies to refining development procedures and to increasing the knowledge base associated with learning techniques highly recommends the continued intertwining of experimentation with development enterprises.

The first set of experiments reported address procedural considerations in the product development process. The second set of studies relate to learning techniques.

A. Development Procedures

1: Effects of Order of Discrimination and Definition Tasks on Pretest Performance

The first study was designed to assess the differential effects of the order in which definition and discrimination tasks were presented on an extended pretest. It was predicted that requiring subjects to formulate definitions of instructional techniques prior to discrimination of their appropriate applications would depress [illegible] tasks. In contrast, subjects not required to overtly formulate definitions or 'rules' to guide their discriminations would be more likely to utilize tentative hypotheses about the defining characteristics and refine these hypotheses as a consequence of a repeated exposure to examples of the technique in the discrimination sequences.

Materials. A preliminary version of the project pretest was the vehicle for the experimentation. The pretest consisted of a section requiring the definition of ten instructional techniques, e.g., prompting, practice. The discrimination section was composed of sample segments of instructional sequences on diverse topics, preceded by specification describing the topic, intended grade level of learners, terminal objective and relevant entry and enroute behaviors. Each sequence included approximately ten frames or two to three pages of material. Two such sequences were provided for each of five instructional techniques. Directions required the subjects to read each section and determine if a given instructional technique was employed in the sequence. Instructional segments were randomized for this portion of the test. In addition, there were three other instructional sequences, preceded by specifications. Subjects were required to engage in multiple discrimination in each of these sequences and to identify which of the five techniques had been employed.

Treatment. The definition section was collated either before (B) or after (A) the discrimination section to construct the two forms of the test relevant to the experimental question.

Dependent Measures. The number of correct discriminations and the total scores served as dependent measures. Twenty-five points were possible on the discrimination section, thirty-five points were possible on the total test.

Subjects. The pretest was administered to two sets of subjects. Twenty-nine UCLA graduate students in education served as one sample. Eleven upper division and graduate students from San Fernando Valley State College participated in a second administration. Considerable variability in performance was anticipated since subjects' experience with the instructional techniques had wide range.

Procedures. For the UCLA study, tests were distributed at random during the first meeting of a class on instruction. Subjects were informed that the task was a pretest and would not affect their grade in the course. Forty-five minutes were allocated for completion of the task.

The attempt to replicate the study at a state college resulted in a number of procedural aberrations that dramatically reduced the number of subjects and control of the experimental situation. For example, that there would be no grading contingency based on pretest results was not made clear to students and sufficient time was not allotted for pretest completion.

Results. Significant differences were not found between means of either subject population on either the discrimination scores or the total scores. The mean discrimination scores of both the UCLA Group (A = [illegible], B = [illegible]) and the state college group (A = [illegible], B = [illegible]) did reveal differences in the predicted direction, however, the respective t values, (t = 1.[illegible], df = 27), (t = .[illegible], df = [illegible]), were not found to be statistically significant. Mean differences between the After and Before groups of approximately the same magnitude were also found on total pretest scores, and these respective t values also were not found to be significant.

Discussion. The results failed to support the prediction that premature formulation of term definitions might depress performance on subsequent discrimination of application of the terms. Further analysis of the experimental treatment suggested that the small number of discrimination tasks per technique, three, may not have been a sufficient number of exposures to exemplars of the technique to promote treatment effects. This interpretation seems an even more plausible possibility when it is considered that for at least three discrimination sequences did not contain an exemplar of the technique at all, i.e., exercise was intended to be a negative instance of the technique. Thus, in light of concept formation literature, two or three exposures to instances of the technique may not have been likely to allow sufficient opportunities to refine hypotheses about the relevant definitional characteristics of the instructional technique. It may be that in testing situations in which the number of discrimination tasks are greater, the relative position of the discrimination task to a definitional task may produce different results.

Although not statistically significant, the results of these studies influenced the organization of the pretest in that, the decision was made to construct the pretest so that the definition task would follow the discrimination task. Superior discrimination of certain techniques implied the reorganization of the project format from a linear to branching style.

II. Effects of Requiring Preparation of Criterion Test Items Before Development of Instruction

A second experiment, involving use of project criterion as a pretest scale, considered the effects of asking developers to prepare test items before they write instructional materials. The lore of product development promotes the practice of generating the criterion test prior to the development of instruction, but to date, no empirical evidence of the efficacy of this procedure has been available. Because the project goal is to enable practioners to produce the most internally valid set of first draft materials possible, the requirement of constructing test items prior to instruction might encourage writers to attend to the critical behaviors necessary to promote their goals.

Subjects. Fourteen graduate students enrolled in a graduate education course participated in the experiment. Seven subjects were randomly assigned to each treatment condition.

Materials. A set of instructional specifications were designed which included the terminal objective (to teach children to tell time to the nearest five minutes) as well as relevant entry and enroute tasks. Directions accompanying the specifications called for the subjects to design an instructional program, taking about sixty minutes of instructional time, to meet the given specifications.

Treatment. The treatment groups differed with respect to whether the contents of the envelope were to be read before or after the subjects wrote the time-telling program. Each envelope contained instructions to construct a criterion test which measured the stated objectives. Directions for opening were inscribed on the front of each envelope.

Dependent Measure. A form was designed to evaluate the extent to which instructional techniques had been incorporated into the program. The evaluation form could be divided into three major categories: the use of instructional techniques, e.g., practice, knowledge of results, the pertinence limits of the instruction, i.e., the observance of specifications and replicability. Three possible points were available for each of nine dimensions and five points were possible for each of two dimensions.

Procedure. Students were provided with specifications accompanied by the treatment-bearing envelopes. They were directed to develop one hour's worth of instruction. For purposes of the course they were to test the results they obtained on one appropriate learner but to make no actual revisions in their first draft.

Analysis and Results. Two scorers independently rated each program without knowledge of treatment conditions. Inter-rater reliability was computed to be .83. Mean ratings for each program were used in the analysis. Results are summarized in Table 30 below:

TABLE 30

Means and Standard Deviations

Criterion Writing Before			Criterion Writing After		
N	$\bar{X}$	s.d.	N	$\bar{X}$	s.d.
7	21.43	6.15	7	15.36	4.41

A significant t value (t = 1.97, 12 degrees of freedom) was obtained suggesting a facilitative effect beyond the .05 level, for writing criterion questions prior to instructional development. Two attempts to replicate the study did not yield a sufficient number of subjects producing complete data to allow analysis of results.

Implications. While the findings of the study reported above case some of the decision process in the particular project under development, and the result of the test-writing study are particularly heartening in their support of product development mythology, the major point of this study should be to encourage those engaged in instructional development to attempt small-scale, controlled variation studies. The findings of such investigations may be limited in impact to the particular project under development and thus be highly 'applied' and delightfully relevant, or possibly have theoretical bases and wider implications. In either cases, the endeavor is worth the attention of development personnel who lack patience for infinite iterations of design-trial and revision.

III: Effect of Promoting a Revision Set on Quantity and Utility of Revision Information

An integral part of product development technology is a commitment to the iterative tryout/revision cycle. A pletora of strategies exist related to the appropriate data to gather and the means to obtain them. The hard-line empirical approach might be first to focus exclusively on performance data gathered from objectively scored tests and reliably operationalized rating scales then to infer revisions from the data sets. However, when materials are in the formative stages, researchers such as Markle (1966) recommend also gathering subject report data. The assumption is that subjects' reactions and suggestions provide a more direct method of identifying instructional and formating deficiencies.

The following studies investigated the effectiveness of informing subjects that they would be asked to suggest program revisions after completing the project materials. We hoped to determine if more extensive judgmental information could be secured from the subjects as a function of directions to attend to the materials as a critic as well as a learner.

A description of each study follows. Implication of the results will be discussed after the description of both studies. A third study was attempted, but a small final sample of five precluded analysis.

Subjects. Eighteen UCLA students enrolled in a secondary teacher preparation course participated in the study. Students volunteered as subjects in exchange for nominal extra credit points in their class.

Materials. First draft versions of text instruction on the techniques of advance organizers, inspection behavior, and task description were presented. Accompanying the text were sets of criterion items for each technique and practice exercises requiring discrimination of the use of the techniques. Subjects were also given an objective and asked to write a short instructional sequence designed to teach that objective.

Treatments. In the "Revision Set" condition an additional page of instructions was colleted with text material. After the title page, a page was inserted which contained the following statement:

> You are participating in this project in order to help the staff revise the materials.
>
> Carefully read the materials presented and try to identify specific things you would change, but do not bother to write them down. You will be asked for this information at the end of the posttest.

The control group received the regularly organized materials.

Dependent Measures. The effects of the treatment were determined by the number of discrete suggestions written by subjects and by the utility of the suggestions as rated on a scale of 1 to 5 by two independent scorers.

Procedure. The experimental materials were randomly distributed to subjects by a member of the project staff. Subjects were informed that the materials would help them improve their design of instructional materials. Subjects worked independently on the materials for approximately three hours. At the end of the session, subjects were requested to indicate on the back of the text and practice exercise sheet any revision suggestions they might have.

Analysis and Results. Two raters independently counted the number of discrete suggestions and rated the utility of each set of suggestions. The inter-rater reliability computed was .82 for the number of suggestions and .84 for the utility ratings. No significant differences were found between the Revision Set and control groups on either of the dependent measures.

Subjects. Twenty-eight students from Arizona State University enrolled in a secondary education course participated. Students were told that they would be using experimental materials, but that there would be no grade attached to their performance on the project tasks.

Materials. In addition to revised versions of the materials presented in the previous study, a second component dealing with Techniques for Response Control (direct practice, knowledge of results and prompting) was included along with its set of criterion items and practice exercises. A pretest, posttest, Revision Sheet and Program Questionnaire were also provided. For the first time, a written set of procedures for use of the materials was prepared to allow the professor of the course to administer the field trial independent of project staff.

Treatment. Half of the materials contained the revision instructions under the title page of the first text section.

Dependent Measures. The number of discrete suggestions and utility of the suggestions were again used as the dependent measures.

Procedures. Materials and directions for their use were administered by the professor of the course, in accordance, instruction was enclosed with the materials. However, the written directions were apparently inadequate, as several long distance phone calls were necessary to clarify the appropriate sequence and manner in which some of the materials were to be used.

Analysis and Results. Data from nineteen of the twenty-eight students participating in the field trial contained revision responses. Raters again tallied the number and utility of revision suggestions independently, yielding inter-rater reliabilities of .85 and .74 respectively, Neither of the measure yielded significant differences between the means of the nine subjects in the Revision Set and the ten subjects in the Control Group.

Discussion. Failure to find statistically significant differences between the Revision Set and Control treatments may be attributable to a number of factors. In the case of the Arizona study, a recurrence of a uniquely-phrased revision recommendation implied a group discussion prior to completion of the Revision Sheet. Therefore, any treatment effects would be expected to be nullified. In addition, the use of a numbered Revision Sheet could have been sufficient to cue equivalent numbers of suggestions. Perhaps the most compelling explanation, applicable to both studies, is that exhorting students to keep revision in mind, without explicitly cueing them to the relevant categories of information, is simply not a powerful enough treatment to produce more information. Results of recent related investigation (Bank, 1972, unpublished doctoral dissertation) suggest that when subjects are directed to attend to the adequacy or experimental materials in terms of communicability (difficulty, sufficiency of information and clarity) and worth (relevance, usefulness and worth) that subsequent revision suggestions relate to these categories. Thus, attempts to elicit more useful revision information from subjects by establishing a Revision Set before they begin instruction should incorporate more specific devices to direct attention to the areas of desired information.

Despite failure to find differences associated with the modest treatment variation, the studies produced valuable revision information relating both to instructional content and organization. In addition, the procedures corroborate the practical value of subject report data, as recommended by Markle (1966).

IV: A Case History of an Instructional Product Development Project with Recommendations for Instructional Technologists

There are few detailed case histories which adequately document the development of research-based, empirically validated instructional products. The study described is a project-related doctoral dissertation, currently in progress, that summarizes, analyzes and interprets the product development procedures utilized and generated by the project. A case history of the course of the project's product development of the study will (a) describe and discuss how the project solved the procedural problems with which it was confronted, and (b) provide recommendations for technologists concerning these procedures and offer suggestions for future research.

The case history of the project will be divided into three parts. The first part will consist of a description of the origins of the product and will examine its primary objectives. Next, the history will focus on the preparation of the project for testing. Finally, the procedural problems of validating the product will be investigated. Two sources of evidence will be used in writing the case history: written records, such as the project proposal and student achievement and affective data, and interviews with the project staff.

Recommendations for instructional technologists concerning the procedures of product development will be derived from a study of the staff's actual and proposed solutions to the problems it confronted. Finally, suggestions for research will be given, based upon the project's procedures and outcomes.

B. Studies Examining the Effects of Learning Techniques in Instructional Materials

V: Effect of Text-embedded Questions on Retention of Text Information

The technique of interspersing questions within text to direct learners' attention to instructionally-relevant content is a procedure that has recently been the subject of a series of investigations (Rothkopf, 1967; Frase, 1970). In a summary of research related to what Rothkopf terms "mathemagenic behavior," attention patterns that facilitate learning, Frase (1970) indicated that a repeated finding of studies is that placing questions after related segments of text increases retention of text content. A set of studies (Glass, 1970) related to the effects of advance organizers or overviews presented before discourse in contrast to summaries presented at the end of instruction indicated that the summary treatment groups obtained superior posttest scores.

It was the purpose of this study to assess the differential effects of questions placed after subsections of text or summaries placed after complete text components.

Subjects. Twenty-four UCLA students enrolled in a graduate course on instruction participated in the experiment. The material was presented as one of the course assignments.

Materials. The text material consisted of five sections dealing with five learning techniques: practice, knowledge of results, prompting, task description, and inspection behavior. The number of pages related to each technique varied from three to seven.

Questions relating to either definitions, characteristics, or conditions for application of each technique were generated for each text section. One question was written for approximately 1-1/2 to 2 pages of text. A total of thirty questions were constructed. The number for each technique ranged from three to five.

Summaries for each technique were composed that included all of the information specifically sought by the questions. The length of the summaries varied from 81 words to 292 words.

Treatment. For the Question (Q) condition, questions were included after a section and set off by broken lines. Typing of text for a technique was continuous, i.e., discourse following a question did not begin on a new page.

The summary for each technique was typed at the end of the text information for that technique. Typing for the Summary condition (S) was also continuous.

Dependent Measures, For each of the five learning techniques a set of criterion items was constructed. Items also related to the definition, characteristics, and conditions under which the technique might vary. The number of items for the technique ranged from six to ten. A total of thirty-eight items were generated.

Procedures. Text information on each of the five learning techniques were presented to the subjects during a regularly scheduled class meeting. Those in the Question group were instructed to record their answers on the materials. Subjects began reading the materials in class and completed them at home. While home use of the materials, implied less experimental control, individual study more closely approximated the intended utilization conditions of the materials.

The criterion test was administered one week later at the beginning of the class meeting.

Analysis and results. Means of the Question and Summary groups were compared on each of the technique scores and on total criterion scores. Tests of the means are presented in Table 31.

TABLE 31

Means, Standard Deviations, and t Values for Questions and Summary Groups on Criterion Items

Topic	Practice (9)[1]		KR (5)		Prompting (6)		Task Description (7)		Inspection Behavior (10)		Total (37)	
	Q	S	Q	S	Q	S	Q	S	Q	S	Q	S
N	13	11	13	11	13	11	13	11	13	11	13	11
X	7.54	7.36	4.54	4.27	5.0	5.0	5.92	5.64	8.46	6.64	31.46	28.91
s.d.	1.13	1.03	.66	1.01	1.00	1.10	.76	.81	1.05	2.29	3.23	4.68
t (df = 22)	.40		.75		0.00		.89		2.43*		1.53	

[1] Number in parentheses indicate the possible score on a measure.

* $p < .05$

A significant t value ($p < .05$) was obtained only on the criterion score for Inspection Behavior. Interestingly, this was the text section describing the value of using embedded questions. The difference between the mean total criterion score of the Question and Summary group favored the Question condition but reached only the .10 significance level.

Discussion. This study again demonstrates the highly dependent nature of the effects of text-embedded questions upon the nature of both the materials and the questions. The text sections of the experimental materials represented discourse which had been repeatedly refined and consolidated to present concise information. Thus, the experimental prose contained a minimum of incidental information as compared with text content usually investigated in text-embedded questions studies.

It should also be noted that the bounds of the information presented in the text allowed generation of questions whose answers primarily contained the most relevant information in the content. The major proportion of the criterion questions tested this same information. This information was also included in the summaries. Thus, the function of the questions and summaries in this study were primarily to transmit direct information, rather than to direct attention to other content not specifically included in the questions or summaries. It would seem that when the structure of the material is "lean," then the necessity of directing attention to categories of relevant information, to help learners separate the wheat from the chaff, is reduced. While the direction of the means on the sub-scores and total criterion score favored the Question condition, further investigation must certainly be conducted to determine the interaction of the effects of text-embedded questions or summaries with the density of instructionally relevant information and uniformity of the prose.

On the basis of the results, both text-embedded questions and summaries were incorporated in the text portions of the project materials.

VI: Effects of Three Characteristics of Text-Embedded Response Requirements on the Development of a Dominant Focus in Prose Learning

Within the context of prose learning, most of the research investigating conditions which control attention patterns, or what Rothkopf (1963) prefers to term inspection behavior, has indicated that the inspection strategies subjects use while reading textual material can be influenced by characteristics of text-embedded questions or directions. A number of studies have indicated that subjects answering within-text questions after reading short prose segments obtained posttest scores superior to pre-question or no-question control groups. In addition, a study by Rothkopf and Bisbicos (1967) partially confirmed their hypothesis that text-embedded questions of restricted categories facilitate learning of restricted categories of text. The interpretation was that inspection behaviors tend to adapt selectively to the nature of text-embedded response requirements (questions).

This study reports the results of a doctoral dissertation related to techniques for promoting effective inspection behaviors. It was one purpose of the study to determine if a mechanism operating in the selective adaptation of inspection behaviors to a restricted class of text-embedded response requirements is a progressive narrowing of focus to restricted categories of content, creating a dominant focus on text information relevant to the class of text-embedded response requirements and a decrease in attention to irrelevant text content. In addition, the study investigated the differential effects of response requirement Type (Question or Objective), Place (Before or After related text), and Order in which two restricted categories of response requirements were presented within the text, (e.g., name questions first, then application questions, or the reverse order).

Method. A 2x2x2 factorial design was employed to investigate treatment effects. The dominant focus hypothesis involved a within-subject treatment of presentation of one restricted category of response requirement within the first eighteen pages of text and a different category of requirement within the last eighteen pages of text.

Forty college students were presented with thirty-six pages of instruction which described six psychological principles. For every three pages, subjects encountered response requirements (either questions or an instructional objective). For one half of the text (pp. 1-18), subjects received response requirements of one restricted category (e.g., identifying appropriate applications of a principle).

A seventy-two item multiple-choice posttest was immediately given, comprising of three kinds of items: items Relevant (R) to the questions of objectives of each three-page text segment; items Non-Relevant (NR) to the response requirement of that segment, but relevant to the other category of response requirement; items Incidental (I) to any segment's response requirement (identifying a date, or name of a teacher of parent which had appeared in that text segment).

Results and Implications. For each restricted category of response requirement (an eighteen-page section of the text), the development of a dominant focus attention pattern was to be inferred if posttest scores on Relevant items were higher than posttest scores on Non-Relevant items. The progressive development of a dominant focus was measured by comparing scores on R items drawn from the first nine pages in which a restricted category of response requirement occurred with the R items drawn from the last nine pages in which the category appeared. The effects of the variables Type, Place, and Order were measured from total posttest scores on R, NR, and I items. A multivariate analysis of variance was employed to assess treatment effects on the three dependent measures.

The prediction that performance on Relevant posttest items would differ significantly from performance on Non-Relevant items was substantiated ($F = 14.54$, $p < .001$, $df = 1,32$). This finding suggested that as a result of exposure to response requirements referring to a limited set of content, a dominant focus does tend to develop on other text content relevant to that restricted category of questions or objectives to the detriment of the attention to Non-Relevant text content. R scores were also significantly greater than NR scores in relation to Type ($F = 13.08$, $p < .001$, $df = 1,32$) and the interaction of Type-Place ($F = 4.626$, $p < .05$, $df = 1,32$) in accordance with other investigators' findings that questions placed after text yield superior posttest performance.

Although the comparison of R scores from the last and the first sections within a response reaquirement category (e.g., pp. 10-18 vs. 1-9) did not reach statistical significance, a graph tracing values of R scores across text segments of the restricted category reveals that the trend of the R scores was to increase from the initial sections in which a category was encountered to the final sections in which the category occurred. While no prediction had been made about the effect of a dominant focus attention pattern upon Incidental items, a comparison of I scores before and after the category switch (pp. 10-18 vs. 19-36) revealed that performance on Incidental items was significantly higher after the switch ($F = 7.31$, $p < .01$, $df = 1,32$), suggesting that the switch yielded a more general search strategy of all text information. After the switch, I scores increased, but R and NR scores decreased slightly.

Comparisons of the effects of Type, Position, and Order upon total R, NR, and I scores again yielded significant values for Type on the R scores ($F = 9.36$, $p < .05$, $df = 1,32$) in favor of questions, and a marginally significant interaction of Type-Place ($p < .05$) in favor of questions after text. These findings suggest that designers and managers of instruction carefully consider the effects which classes of embedded questions may have on the aspects of instruction to which learners attend.

VII: The Effect of Stimulus Variety in Practice Sequences on Discrimination, Application and Attitude Performance

This study, undertaken as a dissertation in the area of instructional product research, focused its investigation on the effects of media in the transmission of practice exercises. While a print version of practice had been formulated, primarily on a cost basis, the study sought to determine if the expected consequences of media presentations, e.g., more active attention, would result in greater performance of subjects. Three treatments were formulated and each consisted of four practice exercises treating four of the five instructional techniques in the project. Use of embedded questions to control text reading behavior patterns was not considered to be appropriate to media variations. One version was rendered in print. A second treatment consisted of the identical material produced in four short, super eight sound motion pictures. The third treatment contained a "variety" condition, one exercise was presented in each of the following media: super eight film, coordinated slide/tape; audio tape only; and print.

Criterion Measures. Discrimination tasks presented in each of the four media were developed. In addition, a transfer task, the ability to write instructional sequences was included as a dependent measure. Finally, an attitude questionnaire was developed in an attempt to gauge the subjects' reaction to the alternative treatment conditions.

Procedures. A pilot study was conducted using nineteen subjects during November, 1971. The purpose of the investigation was to assure that the information was not in the subjects' repertoire and to test the extent to which the criterion measures were appropriate in the procedures.

For the main study, fifty students were randomly assigned to the three treatment groups. Data analysis has not been completed.

CONCLUSIONS

The conclusions one can draw about a development project are necessarily limited. The output of a development effort is a problem solved. The manner of solution and the extent to which the solution pertains to various groups and settings can be explored. The purpose of this section is to describe the outcomes of this project, the procedures developed which seem to have merit, and the problems encountered which others may now anticipate in the future.

Conclusion 1: The materials work.

On the basis of our data, we may conclude that the materials accomplish the primary objective set for them. Participants are able to employ instructional techniques in the writing of first draft programs.

Conclusion 2: The materials present replicable instruction.

We also allege that the sequence of instructional events has been reproduced across validation trials. The final three field tests occurred without intervention by the development staff. Widespread replicability has not been demonstrated, but there is some basis to believe in its existence.

Conclusion 3: A range of persons can learn from the materials.

The materials were tried on preservice education students, graduate students in instructional technology, experienced school personnel, and staff development course writers in an industrial setting. Scores for these groups were in general agreement.

Conclusion 4: Development takes longer than expected.

Even for a staff as confident and well-trained as the one working on this project, the activity of development involves many missteps and cul-de-sacs. A development project does not function like Social Darwinism: there is no linear improvement. The project staff increased in its efficiency as the work progressed, but procedural difficulties occurred enough so that the course of activity was not smooth.

Specific Problems Encountered

Problem 1: Personnel Turnover

Due to a series of untimely events, the continuity of personnel on this project suffered. Part of the problem was related to the phasic need for staff. Heavy use of people during the spring and summer, coupled with the anticipated termination date, led certain of the staff members to depart in favor of employment that offered longer term security. The student status of the staff, while enabling the project to be cost-effective, i.e., employing highly competent people at relatively low wages, also interferred with the project's development. Dissertation and master's thesis work also conflicted with some of the project deadlines.

Problem 2: Production

A consistent difficulty during this project related to the apparently simple problem of getting the materials produced in sufficient quantity on time so that developmental tryouts might be conducted. During the later field trials, a simple set of materials was running about 300 manuscript pages. Thus, for only 35 copies, 10,500 pages required collating. The job increases in magnitude when one considers the time required to retype revised portions of the materials, particularly for those cycles where the turnaround from field trial to field trial was close to only three weeks. The professional staff, with concurrent responsibilities in development and data analysis, was often commandeered to assist in collating packets for field tests. This particular activity was not viewed with much favor by them.

Problem 3: Field Test Sites

Solicitation of sites for developmental and other field testing was described as a time-consuming task. Assistance from other agencies engaged in research and development would have been expected but was not forthcoming. Some procedures should be developed which will promote the acquisition of appropriate field test populations along nationally distributed lines. If an identifiable network of suitable and receptive locations could be supplied, more of the total effort expended would be organized around development rather than site identification activity.

Problem 4: Project Management

After reviewing the products of the project, one might infer that the project management strategy was relatively successful. From the view points of the project manager, however, the project presented difficult management dilemmas. The amount of work was not distributed evenly throughout the period of funding. Alternatives were: (1) to provide regular work for the maximum number of people who would be needed; (2) to provide regular work for the minimum number of people needed, adding individuals to perform specific tasks as necessary. The second option was chosen, primarily because of cost concerns, because of the lack of middle level supervisory personnel (and the budget to sustain them) and because idle staff becomes despondent staff.

A second concern related to student management was related to the appropriate level of supervision and autonomy granted trainees permitted to operate prototypes and use applications which they would employ to guide their development efforts. The emphasis on autonomy perhaps was interpreted as a deficiency in structure. If the project were to be conducted again, it would be recommended that management level personnel operate applications and prototypes for major instructional components, particularly at the outset. The security provided by such models is valuable, even at the cost of perceived creative effort, during the formative stages of a project of this type.

Methodology Developed as a Function of the Project

Sequencing Administration of Criterion Measures

The regular, preferred procedure for sequencing in development work is to administer a pretest, the instructional treatment and a posttest. These tests normally cover the range of competencies included in the instruction. For practical reasons, this project did not adhere to such guidelines. Instead, attention to various testing points was progressively selective throughout the project. An attempt was made to make the time required for testing commensurate with the strength of the instructional package which the trainees was receiving. For instance, pretesting was extensively implanted by subjects who never subsequently received the instruction. The negative consequences of giving over two and one half hours of pretests would be costly to many any positive instructional effects of the treatment. Performance on the pretest averaged about 50% of mastery. Initial product trouble brief forms of the pretest were given, but none which asked the learners to spend an hour approximating the criterion test, the writing of instruction. Similarly, the knowledge posttest section, which measured the ability to produce first-drafts of instruct- was not administered during the relatively early stages of product development. The result felt that subjects would be very unlikely to succeed as a function of reading text and solving discrimination problems. Because it was necessary to husband subject-student time for the most critical purposes, elaborate posttests were sequenced for trouble where the staff had reason to believe that the treatment experience would not [illegible] terminating for the participants.

Testing in Field Tests

The [illegible] plan was most apt to the planning and analysis of field test data. In many pilot or field tests, data from the first should be analyzed primarily for the purpose to verify the validity and performance of the criterion measures. Data in the second field trial should be large enough for the implications to qualitative instructional components. If data from a three field trial is reserved for the purpose of criterion test analyses, one may ultimately be saved the cost of unnecessary revisions made on the basis of faulty criterion data. The usual practice is, of course, to revise texts, locate visual materials, directions,

formats, in other words, the entire package was a performance at each instance of data collection. While such comprehensive strategies may have value, they inhibit the development of a technology of instructional design, in that the reasons for improvement from trial to trial is never explicit. By holding certain of the field trials for the singular purpose of criterion measure evaluation, some progress is made in finding out not only if a product improves, but why.

Progressive Reduction of Staff Influence in Tryouts

Because of the relatively large number of field trials in a short period, this project permitted the staff to sequence the progressive withdrawal of staff involvement in field tests. An important element in the establishment of the exportability of any instructional product is its ability to function without support from the developing agency. While installation or 'implementation' stages of development have been described in long-term, system oriented development projects, they have been unduly ignored in more discrete development missions. By systematically limiting staff involvement, from high intensity (as field trial conductor) through middle levels (as trouble-shooters, observers or telephone counselors) to no involvement, the project staff was able to gain a sense of the power (and limits of power) of the instructional materials. Essentially, field trials themselves were sent through a developmental sequence, it is a step which we would recommend.

Other Outcomes of the Development

The project served to guide three graduate students into the field of instructional development. All are currently employed in development work in other agencies. It supported research for three doctoral dissertations, one of which is completed. It provided needed sustenance for other graduate students who might not have been able to continue in school. Finally, it helped legitimate instructional development as a reasonable, if not fundable, activity for those with university affiliation.

Baker, Eva L. and Schutz, Richard E. "Overview of Four Experiments During the Development of Instructional Materials." Paper presented at the Annual Meeting of the California Educational Research Association, San Diego, April, 1970.

Briggs, L. J. Handbook of Procedures for the Design of Instruction. Pittsburgh: American Institute for Research, 1970.

Legge, Robert M. "Characteristics of Instructional Technologists." NSPI Journal, 1969, Vol. 8, No. 5, 8-9, 14-16.

Legge, Robert M. The Conditions of Learning. New York: Holt, Rinehart and Winston, 1965.

Glaser, Robert. "Psychological Bases for Instructional Design." AV Communication Review, 1967, Vol. 15, 433-449.

Hively, W., Patterson, H. L., and Page, S. H. "A 'Universe-Defined' System of Arithmetic Achievement Tests." Journal of Educational Measurement, 1968, Vol. 5, 275-290.

Lumsdaine, A. A. "Educational Technology, Programmed Learning and Instructional Science." Theories of Learning and Instruction, (E. R. Hilgard, editor), 63rd Yearbook of the National Society for the Study of Education, Part I. Chicago: University of Chicago Press, 1964, 371-401.

Mager, R. F. Preparing Instructional Objectives. Palo Alto, California: Fearon Publishers, 1962.

Markle, S. M. "Empirical Testing of Programs." Programmed Instruction (P. C. Lange, editor), 66th Yearbook of the National Society for the Study of Education, Part II. Chicago: University of Chicago Press, 1967, 104-138.

Markle, S. M. Good Frames and Bad: A Grammar of Frame Writing. (2nd edition) New York: John Wiley and Sons, 1969.

Popham, W. James. "Instructional Product Development: Two Approaches to Training." AV Communication Review, 1967, Vol. 15, 407-411.

Schutz, Richard E. "The Nature of Educational Development." Journal of Research and Development in Education. University of Georgia, 1970, Vol. 1, No. 1, 39-62.

Southwest Educational Development Corporation. "Planning the Systems Approach to Field Testing Educational Products." Calipers. Austin, Texas: Southwest Educational Development Corporation, 1969.

APPENDIX

SELECTIVE ATTENTION

Item Form for Practice Exercises

Single Discrimination

Stimulus Limits

The stimulus must contain three components: directions, specification, and the instructional sequence.

Directions. The following directions are given to the student:

> Read the specifications and instructional sequence which follow. On your answer sheet, circle "yes" or "no" to indicate whether the principle of (insert appropriate principle) has been correctly employed in this sequence of instruction.

Specifications. Preceding the instructional sequence, a description must be provided which specifies:

1. Topic of instructional sequence
2. Approximate grade level for which the instruction is intended
3. Objective(s) governing the instructional sequence

The criterion objective must:

a. Be stated in terms of measurable learner behavior.
b. Represent a cognitive skill which is, preferably, higher than the knowledge level.
c. Include criteria for evaluating the response if a constructed response is required.
d. Be accompanied by one or more enroute objectives if these are also to be dealt with in the instruction.
e. Be accompanied by a statement of any entry behaviors which have been assumed (with the exception of general abilities such as "to read").

Instructional Sequence. The instructional sequence must:

1. Be of approximately 2-5 pages in length
2. Present instruction which is congruent with the objective stated in the specifications, in terms of:

a. Content
b. Response level
c. Scope (covers all objectives specified; does not include large amounts of extraneous information)

Further criteria for the instructional sequences vary depending on the instructional principle being employed for student discrimination. Separate statements of requisite criteria unique to each principle are attached.

Criteria for Evaluating Responses. A response of "yes" is correct when all criteria constituting a correct application of the principle have been met in the instructional sequence. A response of "no" is correct when one or more of the criteria constituting an incorrect application of the principle have been met in the instructional sequence.

Criteria for Instructional Sequences

Inspection Behavior

1. The sequence must consist of textual material (programmed instruction is not appropriate for use with this principle).

2. Within the instruction, sets of questions must be imbedded in one of the following variations:

    a. Correct Application. All of the following criteria must be met to constitute a correct application of the principle.

        --Questions appear after related text segments

        --Questions spaced no more frequently than one set per page of text

        --Questions represent sample items related to the criterion objective or to an enroute objective, in terms of both content

        --Questions clearly require an overt response

    b. Incorrect Application. Any one of the following criteria constitutes an incorrect application:

        --Questions placed before related text segments

        --Questions spaced more frequently than one set per page of text

        --Questions do not represent appropriate sample items for one of the objectives (either content or response level may be incongruent with the objective)

Criteria for Instructional Sequences

Task Description

1. The sequences must consist of textual material (with appropriate imbedded questions) or programmed instruction.

2. Within the instruction, the principle of task description must be employed in one of the following variations.

    a. Correct Application. All of the following criteria must be met to constitute a correct application of princple:

        --Criterion task described at the outset of the instruction (enroute tasks may also be described).

        --Task described in behavioral terms.

        --Task described appropriate sample items congruent with the objective stated in the specifications.

        --Task described in language appropriate for the learner.

    b. Incorrect Application. Any one of the following criteria constitute an incorrect application of the principle:

        --Task described in nonbehavioral (cognitive or affective; statement of intent).

        --Task described does not represent a sample item which is congruent with the objectives stated in the specifications. (Incongruence may reside either in content or response level.)

        --Task is described in language which is inappropriate for the learners (i.e., too complex).

SELECTIVE ATTENTION

Item Form for Practice Exercises

Multiple Discrimination

Stimulus Limits

The stimulus will consist of three components: directions, specifications, and the instructional sequence.

Directions. Students will receive the following directions:

> Read the specifications and instructional sequence which follow. On your answer sheet, circle one or more letter(s) to indicate which instructional principle(s) have been correctly employed in the instruction. Use the following letters:
>
> A. Inspection Behavior
> B. Advance Organizer
> C. Task Description

Specifications. The one-page specification sheet preceding the instructional sequence will conform to those employed in the single discrimination practice exercises for inspection behavior, task description, and advance organizers.

Instructional Sequence. The sequence of instruction shall:

1. Be of approximately 5-7 pages in length.
2. Consist of textual material (not programmed instruction).
3. Be congruent with the specified objectives in terms of content, response level, and scope.
4. Employ the three selective attention principles, as follows:

    a. Embedded questions will be employed (correctly or incorrectly) in each sequence.
    b. At least one of the remaining two principles will be employed (either correctly or incorrectly) in each sequence.
    c. Criteria defining correct and incorrect application of the principles are specified in the item form for single discrimination practice exercises for each principle.

Response Limits

To constitute a correct response, each letter (A,B,C) must be circled if the corresponding principle was correctly employed in the sequence, and must be left uncircled if the corresponding principle was not employed, or was incorrectly employed, in the sequence. A total of three points may be scored for each practice exercise.

TENTATIVE OUTLINE

Staff Development

WORKSHOP ON INSTRUCTIONAL MATERIALS DEVELOPMENT

Investigator's Purpose:	To gather field test data on instructional project.
Time Schedule:	July or August
Duration:	Approximately five hours of instruction, not necessarily on a single day.
Anticipated Audience:	Teachers, Curriculum Developers
Number of Participants:	At least twenty per session
Staff:	Dr. Eva L. Baker, Assistant Professor, UCLA Mrs. Edys Quellmalz, Post Graduate Research Educationist, UCLA Mrs. Judie Safford, Instructional Product Research Specialist Mrs. Adrienne Smith, Research Assistant Mr. Tom McGuire, Research Assistant
Overall Objective:	At the conclusion of the workshop, the participant will be able to produce first draft versions of instructional materials suitable and ready for preliminary field testing for use of students in the age level of the participant's choice.
Enroute Objectives:	To identify and discriminate instances of the use of research-based instructional techniques, including Response Controlling Techniques, such as direct practice, prompting, knowledge of results, task analysis, and Techniques for Directing Attention in written materials, including the use of advance organizers, task describing objectives, and text-embedded questions. To be able to produce specifications for materials, including performance objectives and criteria. To describe the procedure through which materials may be validated.

TENTATIVE WORKSHOP FORMAT

Hour one:	Informal preassessment; purposes of workshop; description of product development procedures.
Hour two:	The writing of performance specification; introduction to materials completion of textual materials and criterion checks for Response Controlling and Attention Directing Techniques.
Hour three:	Practice identifying the use of techniques in single concept instructional sequences.
our four :	Discussion of preceding instruction to identify problems attempts at simulation activities where participants actually develop first draft materials using given instructional techniques.
Hour five:	Critique and exchange of instructional products. Postassessment.

RR 1

Program Questionnaire

Please express your opinions of various aspects of this instruction by circling one number for each item below. Use the following scale:

5 - strongly agree
4 - agree
3 - no opinion
2 - disagree
1 - strongly disagree

Circle only one number for each item.

Expository Sections on Inspection Behavior and Advance Organizers...

1. Were written in clear and understandable terms. . . . 5 4 3 2 1
2. Provided relevant information 5 4 3 2 1
3. Were dull and boring. 5 4 3 2 1
4. Contained superfluous or unnecessary information. . . 5 4 3 2 1
5. Were straightforward and to the point 5 4 3 2 1

Practice Exercises for Each Principle...

6. Used sequences of instruction which were interesting and appropriate 5 4 3 2 1
7. Employed the principles in a way which was ambiguous or too difficult to detect. 5 4 3 2 1
8. Employed the principles in a way which was too obvious and easy to detect. 5 4 3 2 1

In Terms of Length...

9. The package as a whole was too long 5 4 3 2 1
10. The expository sections were too long 5 4 3 2 1
11. Too many practice exercises were provided 5 4 3 2 1

Please provide any specific comments, pro or con, concerning these materials, in whole or in part, by using the reverse side of your answer sheet.

August 17, 1971

CONSTRUCTED RESPONSE CRITERION SHEET

(for use with simulations and criterion performance)

Subject's name _____ Envelope #_____ Scorer_____

Date _____ Location of trial _____

simulation_____ criterion_____ specification #_____

- -

Exclusively text _____ (see section A) Non-text _____ (see section B) Varied_____ (see special instructions B2)

A. Inspection Behavior	Yes	No	
1. Are questions embedded in text?	___	___	(1)
If yes,			
a. Do questions appear following relevant text?	___	___	(2)
b. Are questions related to objectives stated in specifications?	___	___	(2)
c. To how many terminal and/or enrbute objectives are questions explicitly related?	(circle) 1	2	3 4
Add value of each yes plus number circled in "3"			
	subscore_____		

B. Task Description	Yes	No	
1. Is an operational* task description provided?	___	___	(2)
If yes,			
a. Is language appropriate to learner's level?	___	___	(1)
b. Is task description provided prior to instruction?	___	___	(2)
c. For how many terminal or enroute objectives are task descriptions provided?	(circle) 1	2	3 4
Add value of each yes plus number circled in "c"			
	subscore_____		

*operation means content and behavior are described which conform to either terminal or enroute objectives included in specifications.

B. - 2 not to be completed if A-1 is "yes"

Direct Practice	Yes	No	
2. Is practice explicitly provided relevant to the objectives?	___	___	(1)
If yes,			
a. Was more than a single practice opportunity provided for any objective?	___	___	(1)
b. Was practice provided for multiple content examples or in a variety of contexts?	___	___	(2)
c. For how many terminal or enroute objectives was practice provided?	(circle)	1	2 3 4
Add value of each yes plus number circled in "c"			
	subscore ___		

Knowledge of Results	Yes	No	
3. Was knowledge of results provided for either overt or covert responses?	___	___	(1)
If yes,			
	effective (1)	Not effective (0)	
a. Estimate effectiveness in terms of immediacy.	___	___	
	Yes	No	
b. Form of knowledge of results is appropriate to response requirements	___	___	(1)
c. Estimate the proportion of responses which receive knowledge of results	(circle)	4	3 2 1
		100%	80% 60% 40%
Add value of yes plus number circled in "c"			
	subscore ___		

3. - 4 Prompting

		Yes	No	
a.	Were prompts used?	___	___	(1)

If yes,

a.	Did the writer tend to avoid copy frames or give-aways as primary prompting techniques?	___	___	(1)
b.	Were prompts primarily substantive rather than formal?	___	___	(1)

c. Were prompts faded, i.e., were unprompted responses called for during latter phases of instruction? (circle) gradual fading 4 3 2 1 abrupt/no ___

Add value of each yes plus number circled in "c"

subscore_____

Technique total score_____

	Yes	No	
Language appropriate to learner	___	___	(2)
Replicable	___	___	(2)

Total _____

Revision August 23, 1971

CONSTRUCTED RESPONSE CRITERION SHEET

(for use with simulations and criterion performance)

Subject's name_____ Envelope #_____ Scorer_____

Date of trial _____ Location of trial_____

simulation_____ criterion_____ specification #_____

- -

Exclusively text_____ (Omit section C)
Non-text_____ (Omit section A)
Varied_____ (Omit section D if A is used)

A. Inspection Behavior	Yes	No	
Are questions embedded in text? If yes,	____	____	(2)
1. Do questions appear following relevant text?	____	____	(2)
2. Are questions related to criterion objective stated in specifications?	____	____	(2)
3. To how many enroute objectives are questions explicitly related?	(circle) 1 2 3		
Add value of each yes plus number circled in "3"	Subscore_____ (maximum = 9)		

B. Task Description	Yes	No	
Is an operational task description provided? If yes,	____	____	(2)
1. Is language appropriate to learner's level?	____	____	(1)
2. Is task description provided prior to instruction?	____	____	(2)
3. Is task description provided for the criterion objective?	____	____	(2)
4. For how many enroute objectives are task descriptions provided?	(circle) 1 2 3		
Add value of each yes plus number circled in "4"	Subscore_____ (maximum = 10)		

STAFF UTILIZATION

The following individuals were employed by the project:

Sheila Allen

Britta Bull

Barbara Donner

Arlene Fink

Sherry Frankel

Michael Kosecoff

Tom McGuire

Linda Morishita*

Edys Quellmalz*

Judie Safford

Aleta Osborne Saloutos

Karen Schwartz

Gayla Sexy

Adrienne Smith

Howard Sullivan

Lee Trithart

Lena Wackenstedt

In addition, Lani Steele and Barbara Bosak worked free of charge on materials related to the project; Arlene Fink provided unpaid assistance.

*more than six months